

FIGURE 1

CCAGGTCCA ACTGCACCTCGGTTCTATCGATTGAATTCCCCGGGGATCCTCTAGAGATCCCT
CGACCTCGACCCACCGTCCGCCAAGCTGGCCCTGCACGGCTGCAAGGGAGGCTCCTGTGGA
CAGGCCAGGCAGGTGGGCCTCAGGAGGTGCCTCCAGGCCAGTGGCCTGAGGCCAGC
AAGGGCTAGGGTCCATCTCCAGTCCCAGGACACAGCAGCGGCCACCATGCCAAGCCTGGC
TCCAGCAGCATCAGCAGCCCCCAGGACCGGGGAGGCACAGGTGGCCCCACCACCCGGAGG
AGCAGCTCCTGCCCTGTCCGGGGA**TG**ACTGATTCTCCTCGCCAGGCCACCCAGAGGAGA
AGGCCACCCCGCCTGGAGGCACAGGCCATGAGGGCTCTCAGGAGGTGCTGCTGATGTGGCT
TCTGGTGGCAGTGGCGCACAGAGCACGCCTACCGGCCGGCGTTAGGGTGTGCT
GTCCCAGGCTCACGGGACCCCTGTCTCGAGTCGTTCGTCAGCGTGTGACCAGCCTCC
TCACCACCTGCGACGGCACCGGCCCTGCAGCACCTACCGAACATTATAGGACCGCCTAC
CGCCGCAGCCCTGGCCTGCCAGGCCTCGCTACGCGTGTGCCCGGCTGGAAGAG
GACCAGCGGGCTTCCTGGGCCTGTGGAGCAGCAATATGCCAGCCGCATGCCGGAACGGAG
GGAGCTGTGTCCAGCCTGGCGCTGCCGCTGCCCTGCAGGATGGCGGGGTGACACTGCCAG
TCAGATGTGGATGAATGCAGTGCTAGGAGGGCGGCTGTCCCCAGCGCTGCATCAACACCGC
CGGCAGTTACTGGTGCAGTGTGGAGGGCACAGCCTGTCTGCAGACGGTACACTGTG
TCCCCAAGGGAGGGCCCCCAGGTGGCCCCAACCCGACAGGAGTGGACAGTGCAATGAAG
GAAGAAGTGCAGAGGCTGCAGTCCAGGGTGGACCTGCTGGAGGAGAACGCTGCAGCTGGT
GGCCCCACTGCACAGCCTGGCCTCGCAGGCACTGGAGCATGGCTCCGGACCCGGCAGCC
TCCTGGTGCACTCCTCCAGCAGCTGGCCGCATCGACTCCCTGAGCGAGCAGATTCCCTC
CTGGAGGAGCAGCTGGGTCTGCTCTGCAAGAAAGACTCGTGATGTGCCAGCGCCCCAGG
CTGGACTGAGCCCTCACGCCGCCCTGCAGCCCCATGCCCTGCCAACATGCTGGGGTC
CAGAACCCACCTCGGGGTGACTGAGCGGAAGGCCAGGCAGGGCTTCCTCCTTTCTCCTC
CCCTCCCTCGGGAGGGTCCCCAGACCCCTGGCATGGGATGGCTGGGATTTTTGTGAAT
CCACCCCTGGCTACCCCCACCCCTGGTTACCCAACGGCATCCAAAGGCCAGGTGGCCCTCA
GCTGAGGGAGGTACGAGTTCCCTGCTGGAGCCTGGGACCCATGGCACAGGCCAGGCAGCC
CGGAGGCTGGGTGGGCCTCAGTGGGGCTGCTGCCTGACCCCCAGCACAATAAAATGAAA
CGTGA
AGAGTCGACCTGCAGAAGCTGGCCCATGCCAACCTGTTATTGCAGCTTATAATGGT
TACAAAT

FIGURE 2

MTDSPPPGHPEEKATPPGGTGHEGLSGGAADVVASGVGSGRHRARLPARPLGCVLSRAHDPV
SESFVQRVYQPFLTCDGHRACTYRTIYRTAYRRSPGLAPARPRYACCPGWKRTSGLPGAC
GAAICQPPCRNGGSCVQPGRCRCPAGWRGDTCQSDVDECSARRGGCPQRCINTAGSYWCQCW
EGHSLSADGTLCKVPKGPPRVAPNPTGVDSAMKEEVQRLQSRVLLEEKLQLVLAPLHSLAS
QALEHGLPDPGSLLVHSFQQQLGRIDSLSEQISFLEEQLGSCSCKKDS

Signal sequence:

amino acids 1-19

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 93-97, 270-274

N-myristoylation sites.

amino acids 19-25, 78-84, 97-103, 100-106, 103-109, 157-163,
191-197, 265-271

Amidation site.

amino acids 26-30

Aspartic acid and asparagine hydroxylation site.

amino acids 152-164

Cell attachment sequence.

amino acids 130-133

EGF-like domain cysteine pattern signature.

amino acids 123-135

FIGURE 3

CGCTCGCCCCGTGCCCTCGCCTCCCCGAGAGTCCCCCGCAGCAGATGTGTGGG
GTCAGCCCACGGGGGACTATGGTGAATCCCGCGCTACGCACACTGGCCCTGATC
CGGTTCTGGTGCCTGGCATACCAACATAGCCATCGACTCGGGGAGCAGGCCTGAA
CCGGGGCATTGCTGCTCAAGGAGGATGCAGTCGAGATGCTGGCCAGCTACGGGCTGGCGT
ACTCCCTCATGAAGTTCTCACGGGTCCATGAGTGAATTCAAAAATGTGGGCTGGTGT
GTGAACAGCAAGAGAGACAGGACAAAGCCGTCTGTGTAGGTGGCAGGGGCATCGC
TGCCGTCTTCACACACTGATAGCTTATAGTGATTAGGATACTACATTATCAATAACTGC
ACCATGTGGACGAGTCGGTGGGAGCAAGACGAGAAGGGCCTCCTGTACCTCGCCGCTT
CCTTCATGGACGCAATGGCATGGACCCATGCTGGCATTCTTAAACACAAATACAGTT
CCTGGTGGGATGTGCCTCAATCTCAGATGTCATAGCTCAGGTTGTTTGAGCCATTG
TTCACAGTCACCTGGAATGCCGGAGCCCTGCTCATCCGATCCTCCTGTACATGGC
GCACTGTGCGCTGCACCACCCCTGTGCCTGGCTACTACAAGAACATTACGACATCATCCC
TGACAGAAGTGGCCCGGAGCTGGGGGAGATGCAACAATAAGAAAGATGCTGAGCTCTGGT
GGCCTTGGCTCTAATTCTGCCACACAGAGAATCAGTCGGCTATTGTCAACCTTTGTT
TCCCAGGACCTTGGCAGTTCTGCAGCCACAGAGGAGTGGCATTGACAGCCACATA
CCCTGTGGGTACATGCCATACGGCTGGTACGGAAATCCGTGCTGTATCCTGCTTCG
ACAAGAATAACCCAGCAACAAACTGGTGAGCACGAGCAACACAGTCACGGCAGCCCACATC
AAGAAGTTCACCTCGTCTGCATGGCTCTGCACTCACGCTCTGTTGATGTTTGAC
ACCCAACGTGCTGAGAAATCTTGATAGACATCATGGAGTGGACTTGCCTTGCAGAAC
TCTGTGTTGTTCTTGCGATCTCTCCTTCCAGTCCAGTCAGTCACAGTGAGGGCGCAT
CTCACCGGGTGGCTGATGACACTGAAGAAAACCTTCGTCCTGCCCGAGCTCTGCTGCG
GATCATCGCCTCATGCCAGCCTGTTCTACCCCTACCTGGGGTGCACGGTGCACCC
TGGCGTGGCTCCCTCTGGCGGGCTTGTGGAGAATCCACCATGGTCGCCATGCTGCG
TGCTATGTCACCGGAAGCAGAAAAAGAAGATGGAGAATGAGTCGGCACGGAGGGGAAGA
CTCTGCCATGACAGACATGCCCTCGACAGAGGAGGTGACAGACATCGTGGAAATGAGAGAGG
AGAATGAATAAGGCACGGGACGCCATGGCACTGCAGGGACGGTCAGTCAGGATGACACTTC
GGCATCATCTCTCCCTCTCCATCGTATTTGTTCCCTTTTTGTTGTTGGTAAT
GAAAGAGGCCTGATTTAAAGGTTCTGTCAATTCTCTAGCATACTGGTATGCTCACACT
GACGGGGGGACCTAGTGAATGGTCTTACTGTTGCTATGAAAAACAAACGAAACAACGAC
TTCATACCCCTGCCTCACGAAAACCAAAAGACACAGCTGCCTCACGGTTGACGTTGTG
TCCCTCCCTGGACAATCTCCTCTTGAACCAAAGGACTGCAGCTGTGCCATCGCCTCG
CACCCCTGCACAGCAGGCCACAGACTCTCCTGTCCTCATCGCTCTTAAGAATCAACAGG
TTAAAACCGGCTTCTTGTGATTGCTTCCAGTCACATGGCGTACAAAGAGATGGAGCCC
CGGTGGCCTCTTAAATTCCCTCTGCCACGGAGTTGAAACCATCTACTCCACACATGCAG
GAGGCAGGGTGGCACGCTGCAGCCCCGAGTCCCCGTTCACACTGAGGAACGGAGACCTGTGAC
CACAGCAGGCTGACAGATGGACAGAAATCTCCGTTAGAAAGGTTGGTTGAAATGCCCGGG
GGCAGCAAACGTGACATGGTGAATGATAGCATTCTCGTTCTCTAGATCTGAGCAA
GCTGTCAGTTCTCACCCCCACCGTGTATATACATGAGCTAACCTTTAAATTGTACAAAAA
GCGCATCTCCAGATTCCAGACCCCTGCCGATGACTTTCTGAAGGCTGCTTCCCTCG
CTTCCTGAAGGTGCGATTAGAGCGAGTCACATGGAGCATCTTAACCTTGATTTAGTTT
TACAGTGAACTGAAGCTTAAGTCTCATCCAGCATTCTAATGCCAGGGTGTAGGGTAAC
TTTGAAGTAGATATATTACCTGGTTCTGCTATCCTAGTCATAACTCTCGGGTACAGGTA
TTGAGAATGTAATCGGTACTTCCCTCCCACACCATACGATAAAAGCAAGACATTATAACG
ATACCAGAGTCACTATGTGGTCTCCCTGAAATAACGCATTGAAATCCATGCAGTCAGTA
TATTTCTAAGTTGGAAAGCAGGTTTCTTAAAAAATTATAGACACGGTTCACT
AAATTGATTAGTCAGAATTCTAGACTGAAAGAACCTAAACAAAAAAATTAAAGATA
TAAATATATGCTGTATGTATGTAATTATTTAGGCTATAATACATTTCCTATTTCGC
ATTTCATAAAATGTCTCTAATACAAAAAA

FIGURE 4

MVKFPALTHYWPLIRFLVPLGITNIAIDFGEQALNRGIAAVKEDAVEMLASYGLASLMKFF
TGPMMSDFKNVGLVFVNSKRDRTKAVLCMVVAGAIAAVFHTLIAYSDLGYYIINKLHHVDESV
GSKTRRAFLYLAAPFMDAMAWTAGILLKHYSFLVGCAISDVIAQVVFVAILLHSHLEC
REPLLIPILSLYMGALVRCTTLCGYYKNIHDIIPDRSGPELGGDATIRKMLSFWWPLALIL
ATQRISRPIVNLFVSRLGGSSAATEAVAILTATYPVGHMPYGLTEIRAVYPAFDKNNPSN
KLVSTSNTVTAAHIKKFTFVCMALSLTLCFVMFWTPNVSEKILIDIIGVDFAFELCVVPLR
IFSFFPVPTVRAHLTGWLMTLKKTFLAPSSVLRIVLIALSVLPYLGVHGATLGVGSL
AGFVGESTMVAIAACYVYRKQKKKMENESATEGEDSAMTDMPPTEEVTDIVEMREENE

Transmembrane domains:

amino acids 86-106, 163-179, 191-205, 237-253, 327-343, 357-374,
408-423, 431-445

50
40
30
20
10
0

FIGURE 5

CCTGACAGAAGTCCCCGGAGCTGGGGAGATNCAACATTAAGAAGATGCTGAGCTCTGGT
GCCNTTGGCTCTAATTCTGGCCACACAGAGAANCAGTCGGCTATTGTCAACCTCTTGTT
TCCCAGGACCTTGGTGGCAGTTCTGCAGCCACAGAGGCAGTGGCGATTTGACAGCCACATA
CCCTGTGGGTACATGCCATACGGCTGGTTGACGGAAATCCGTGCTGTATCCTGCTTCG
ACAAGAATAACCCCAGCAACAAACTGGTGAGCAGCAACACAGTCACGGCGGCCACATC
AAGAAGTTCACCTCGTCTGCATGGCTCTGTCACTCACGCTCTGTTCGTATGTTGGAC
ACCCAACGTGTCTGNGAAATCTTGATAGACATCGGAGTGGACTTTGCCTTGAGAAC
TCTGTGTTGTTCTTGCAGATCTTCTCCTTCTTCCAGTTCCAGTCACAGTGAGGGCGCAT
CTCACCGGGTGGCTGATGACACTGAAGAAAACCTTCGTC

FIGURE 6

TGACGGAATCCGGGCTGGTATCCTGGTTNGACAAGATAAACCCCCAGCAANAAATTGGG
GAGCAGGGCAAAACAGTNACGGGCAGCCCACATCAAGAAGTTCACCTNGTTGNATGGNTC
TGTCAACTCACGCTNTGTTCGTATGTTGGACACCCAAAGTGGTGGAGAAAATTTGAT
AGACATNATCGGAGTGGANTTGCCTTGAGAANTTGNNGNTGTTCCGGATTTCT
CCTTTTCCCAGTCCAGTCACAGNGAGGGCGCATCTCACCGGGNGNTGATGACANTGAAG
AAAACCTTGTCTTGCCTTGCCTTGCAGCTNTTGGTGCAGTCATTGTCCTNATNGCCAGCCTG
GGTCCTACCTACCTGGGGTGCACGGTGCAGCCCTGGCGTGGTCCCTGGCGGGCA

FIGURE 7

TATTCCCAGTTCCGGTCACGGGGAGGGCGCATNTCACCGGGTGGCTGANGACACTGAAGAAA
ACCTTNGTCTTGCCCCAGNTTGTGNTGCGGATNATCGTCCTCATGCCAGCCTNGTGGT
CCTACCCTACCTGGGGTGCACGGTGAGAC

30
20
10
0

FIGURE 8

CCCCCGCGCCCGGCCGGCGCCGAAGCCGGAGCCACCGCC**ATGGGGCCTGCCTGG**
GCCTGCTCCCTGCTCAGCTGCGCGCTGCCTCTGCGGCTCTGCCCTGCATCCTGTGCAG
CTGCTGCCCGCCAGCGCAACTCCACCGTGAGCCGCCTCATCTTCACGTTCTCCTCTTCC
TGGGGGTGCTGGTGTCCATCATTATGCTGAGCCCAGGGCTGGAGAGTCAGCTCTACAAGCTG
CCCTGGGTGTGAGGAGGGGCCGGATCCCCACCGCCTGCAGGCCACATGACTGTGG
CTCCCTGCTGGCTACCGCGCTGTCTACCGCATGTGCTTCGCCACGGCGGCCCTCTTCT
TCTTTTCACCCGCTCATGCTCTGCGTGAGCAGCAGCCGGACCCCCGGCTGCCATCCAG
AATGGGTTTGGTTTTAAGTTCTGATCCTGGTGGCCTCACCGTGGTGCCTTACAT
CCCTGACGGCTCCTCACCAACATCTGGTTCTACTTCGGCGTGTGGCTCCTCCTTCA
TCCTCATCCAGCTGGTGTGCTCATCGACTTGCACACTCCTGGAACCAGCGGTGGCTGGC
AAGGCCGAGGAGTGCATTCCGTGCCTGGTACGCAGGCCTCTTCTTCACTCTCCTT
CTACTTGCTGTCGATCGCGCCGTGGCGCTGATGTTCATGTACTACACTGAGGCCAGCGGCT
GCCACGAGGGCAAGGTCTTCATCAGCCTAACCTCACCTCTGTGTCTGCGTGTCCATCGCT
GCTGTCCTGCCAAGGTCCAGGACGCCAGCCAACTCGGGCTGCTGCAGGCCTCGGTCT
CACCCCTACACCATGTTGTCACCTGGTCAGCCCTATCCAGTATCCCTGAACAGAAATGCA
ACCCCCATTGCCAACCCAGCTGGCAACGAGACAGTTGTGGCAGGCCCGAGGGCTATGAG
ACCCAGTGGTGGATGCCCGAGCATTGTGGCCTCATCATCTCCTCTGTGCACCCTT
CATCAGTCTGCGCTCCTCAGACCACCGCAGGTGAACAGCCTGATGCAGACCGAGGAGTGCC
CACCTATGCTAGACGCCACACAGCAGCAGCAGCAGCAGGTGGCAGCCTGTGAGGGCCGGC
TTTGACAACGAGCAGGACGGCGTCACCTACAGCTACTCCTCTTCACTTCTGCCTGGTCT
GCCCTCACTGCACGTCATGATGACGCTACCAACTGGTACAAGCCGGTGAGACCCGGAAGA
TGATCAGCACGTGGACCGCCGTGTGGTGAAGATCTGTGCCAGCTGGCAGGGCTGCTCCTC
TACCTGTGGACCCTGGTAGCCCCACTCCTCTGCGCAACCGCGACTTCAGC**TGAGGCAGCCT**
CACAGCCTGCCATCTGGTGCCTCCTGCCACCTGGTGCCTCTGGCTCGGTGACAGCCAACCT
GCCCTCCCCACACCAATGCCAGGCTGAGCCCCACCCCTGCCAGCTCCAGGACCTG
CCCTGAGCCGGGCTTCTAGTCGTAGTGCCTCAGGGTCCGAGGAGCATCAGGCTCCTGCA
GAGCCCCATCCCCCGCCACACCCACACGGTGGAGCTGCCTTCCCTCCCTCCCTGT
TGCCCATACTCAGCATCTGGATGAAAGGGCTCCCTGTCTCAGGCTCCACGGGAGCGGG
CTGCTGGAGAGAGCGGGAACTCCCACACAGTGGGCATCCGGCACTGAAGCCCTGGTCT
CCTGGTCACGTCCCCCAGGGGACCCCTGCCCTGGACTTCGTGCCTTACTGAGTCTCT
AAGACTTTTCTAATAACAAGCCAGTGCCTGTAAAAAAA

FIGURE 9

MGACLGACSLSCASCLCGSAPCILCSCCPASRNSTVSRLIFTFFLFLGVLVSIIMLSPGVE
SQLYKLPWVCEEGAGIPTVLQGHIDCGSLLGYRAVYRMCATAAFFFFFFTLLMLCVSSSRD
PRAAIQNGFWFFKFLILVGLTVGAFYIPDGSFTNIWFYFGVVGSFLFILIQLVLLIDFAH
NQRWLGAEECDSRAYAGLFFFTLLFYLLSIAAVALMFYYTEPSGCHEGKVFISLNLTFC
VCVSIAAVLPKVQDAQPNSGLLQASVITLYTMFVTWSALSSIPEQKCNPHLPTQLGNETVVA
GPEGYETQWWDAPSIVGLIIFLLCTLFISLRSSDHRQVNNSLMQTEECPPMLDATQQQQQQVA
ACEGRAFDNEQDGVTYSYFFFHFCVLASLHVMMTLTNWYKPGTRKMISTWTAVWVKICAS
WAGLLLWTLVAPLLLNRNDFS

Signal sequence:

amino acids 1-20

Transmembrane domains:

amino acids 40-58, 101-116, 134-150, 162-178, 206-223, 240-257,
272-283, 324-340, 391-406, 428-444

FIGURE 10

GAGCGAGGCCGGGACTGAAGGTGTGGTGTGAGCCCTTGGCAGAGGGTTAACCTGGTC
AAATGCACGGATTCTCACCTCGTACAGTTACGCTCTCCGCGGACGTCCGAGGGACTTGA
AGTCCTGAGCGCTCAAGTTGTCCGTAGGTCGAGAGAAGGCCATGGAGGTGCCGCCACCGC
ACCGCGGAGCTTCTGTAGAGCATTGTGCCTATTCCCCGAGTCTTGCTGCCAGCTG
TGACTGCCGATTCGGAAGTCCTTGAGGAGCGTCAGAACGGCTCCCTACGTCCAGAGCCC
TATTACCGGAATCTGGATGGACCGCCTCCGGAGCTGTTGGCAAAGATGAACAGCAGAG
AATTCAAAGGACCTTGCTAATATCTGTAAGACGGCAGCTACAGCAGGCATCATTGGCTGG
TGTATGGGGAATACCAGCTTATTGCTAAACAATACATTGAGCAGAGCCAGGCA
GAAATTATCATAACCGGTTGATGCTGCAATCTGCACATCGTGTGCCACAGGAGCTT
CATTGTTATGGCTGGCGCTGGGTTGGAGAACTGCAGTGTGACTATATTCAACACAG
TGAACACTAGTCTGAATGTATACGAAATAAGATGCCATTGTAATTGCAGGA
GCTGTCACGGAAAGTCTTTAGGATAAACGTTAGGCCTGCGTGGCCTGGTGGCTGGCAT
AATTGGAGCCTTGTGGCACTCCTGTAGGAGGCCTGCTGATGGCATTCAAGTACGCTG
GTGAGACTGTTAGGAAAGAAAACAGAAGGATCGAAAGGCACTCCATGAGCTAAACTGGAA
GAGTGGAAAGGCAGACTACAAGTTACTGAGCACCTCCCTGAGAAAATTGAAAGTAGTTACG
GGAAGATGAAACCTGAGAATGATGCTAAGAAAATTGAAAGCACTGCTAACCTTCCTAGAAACC
CTTCAGTAATAGATAAACAGACAAGGACTTGAAAGTGCTCTGAAACTCAGGAGA
GCTGAAGGGAGCTGCCATGTCCGATGAATGCCAACAGACAGGCCACTCTTGGTCAGCCTGC
TGACAAATTAAAGTGTGGTACCTGTGGTGGCAGTGGCTTGCTTTGTCTTTCTTCTT
TTAACTAAGAATGGGCTGTTGACTCTCACTTACTTATCCTAAATTAAACATAACT
TATGTTGTATTAATCTATCAATATGCATACATGGATATATCCACCCACCTAGATTTAA
GCAGTAAATAAACATTGCAAAAGATTAAAGTTGAATTACAGTT

FIGURE 11

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA23318
><subunit 1 of 1, 285 aa, 1 stop
><MW: 32190, pI: 9.03, NX(S/T): 2
MEVPPPAPRSFLCRALCLFPRVFAAEAVTADSEVLEERQKRLPYVPEPYYPESGWDRRLRELF
GKDEQQRISKDLANICKTAATAGIIGWVYGGIPAFIHKQQYIEQSQAEIYHNRFDQVSAH
RAATRGFIRYGRWGRWRTAVFVTIFNTVNTSLNVYRNKDALSHFVIAGAVTGSFRINVGLR
GLVAGGIIGALLGTPVGGLLMAFQKYAGETVQERKQKDRKALHELKLEEWKGRLQVTEHLPE
KIESSLREDEPENDAKKIEALLNLPRNPSVIDKQDKD
```

Important Features:

Signal Peptide:

amino acids 1-24

Transmembrane domains:

amino acids 76-96 and 171-195

N-glycosylation site:

amino acids 153-156

FIGURE 12

CGGAAGTCCCTTGAGGAGCGTCAGAAGCGGCTTCCCTACGTCCAGAGCCCTATTACCGGA
ATCTGGATGGGACCGCTCCGGGAGCTGTTGGCAAAGATGAACAGCAGAGAATTCAAAGGA
CCTTGCTAATATCTGTAAGACGGCAGCTACAGCAGGCATCATTGGCTGGGTGTATGGGGAA
TACCAAGCTTTATTCATGCTAAACAACAATACATTGAGCAGAGCCAGGCAGAAATTATCAT
AACCGGTTGATGCTGTGCAATCTGCACATCGTGCTGCCACACGAGGCTTCATTGTTCATG
GCTGGCGCCGAACC

FIGURE 13

TCAAGTTGTCCGTAGGTCGAGAGAAGGCCATGGAGGTGCCGCCACCGGCACCGCGGAGCTT
TTTCTGTAGAGCATTGTGCCATTCCCCGAGTTTGCTGCCGAAGCTGTGACTGCCGAT
TCGGAAGTCCTTGAGGAGCGTCAGAACGGCTCCCTACGTCCCAGAGCCCTATTACCGGA
ATTGGATGGGACCGCCTCCGGAGCTGTTGGCAAAGATGAACAGCAGAGAATTCAAAGG
ACCTTGCTGATATNTGTAAGACGGCAGCTACAGCAGGCATCATTGGCTGGGTGTATGGGGGA
ATACCAGCTTTATTGNTAAACAACAATACATTGAGCAGAGCCAGGCAGAAATTATNA
TAACC

FIGURE 14

GAGCCGCCGCCGCGCGCGCGCGCACTGCAGCCCCAGGCCCGGCCACCCACGTCT
GCGTTGCTGCCCGCCTGGCCAGGCCCCAAAGGCAAGGACAAAGCAGCTGTCAAGGAACCT
CCGCCGGAGTCGAATTCACGTGCAGCTGCCGCAACCACAGGTTCAAG**ATGGTTGCGGGG**
GCTTCGCGTGTCCAAGAACTGCCTGTGCCCTAACCTGCTTACACCTGGTTAGTCTG
CTGCTAATTGGAATTGCTGCGTGGGCATTGGCTCAGGCTGATTCCAGTCTCGAGTGGT
CGCGTGGTCATTGCAGTGGCATCTTCTGTTAGCTGATTGCTTAGTGGTCTGATTGGAG
CTGTAAAACATCATCAGGTGTTGCTATTTTATATGATTATTCTGTTACTTGTATTATT
GTTCAAGTTCTGTATCTGCGCTTGTAGCCCTGAACCAGGAGCAACAGGGTCAGCTCT
GGAGGTTGGTGGAACAAACGGCAAGTGCTCGAAATGACATCCAGAGAAATCTAAACTGCT
GTGGGTTCCGAAGTGTAAACCAAATGACACCTGTCTGGTAGCTGTGTTAAAGTGACCAC
TCGTGCTGCCATGTGCTCCAATCATAGGAGAATATGCTGGAGAGGTTGAGATTGTTGG
TGGCATTGGCCTGTTCTCAGTTTACAGAGATCCTGGGTGTTGGCTGACCTACAGATACA
GGAACCAGAAAGACCCCCGCGCGAACCTCTAGTGCATTCTT**TGATGAGAAAACAAGGAAGAT**
TTCCTTCGTATTATGATCTGTTCACTTCTGTAATTTCTGTTAAGCTCCATTGCCAGT
TTAAGGAAGGAAACACTATCTGGAAAAGTACCTTATTGATAGTGGAAATTATATTTTACT
CTATGTTCTCTACATGTTTTCTTCCGTTGCTGAAAAATATTGAAACTTGTGGTCTC
TGAAGCTCGGTGGCACCTGGAATTACTGTATTCACTGTCTGGCACTGTCCACTGTGGCCTT
TCTTAGCATTTCACCTGCAGAAAAACTTGTATGGTACCACTGTGTTGGTATATGGTAA
TCTGAACGTACATCTCACTGGTATAATTATGTAGCACTGTGCTGTAGATAGTCCTAC
TGGAAAAAGAGTGGAAATTATTAAATCAGAAAGTATGAGATCCTGTTATGTTAAGGAAA
TCCAAATTCCAATTGGTCTTTAGGAAAGATTGTTGGTAAAAAGTGTAGTA
AAAAATGATAATTACTGTAGTCTTTATGATTACACCAATGTATTCTAGAAATAGTTAT
GTCTAGGAAATTGTGGTTAATTGACTTTACAGGTAAGTGCAAAGGAGAAGTGGTT
CATGAAATGTTCTAATGTATAATAACATTACCTCAGCCTCCATCAGAATGGAACGAGTT
TGAGTAATCAGGAAGTATCTATATGATCTGATATTGTTATAATAATTGAAAGTCTAA
AAGACTGCATTAAACAAGTTAGTATTAAATGCGTGGCCACGTAGCAAAAGATATTG
ATTATCTAAAAATTGTTAAATACCGTTCTGAAATTCTCAGTATTGAAACAGCAACTT
GTCACACCTAACGATATTGAATATGATCTCCATAATTGAAATTGAAATCGTATTGTG
GCTCTGTATATTCTGTTAAAAATTAAAGGACAGAAACCTTCTTGTATGCATGTTGA
ATTAAAAGAAAGTAATGGAAG

FIGURE 15

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA39979

><subunit 1 of 1, 204 aa, 1 stop

><MW: 22147, pI: 8.37, NX(S/T): 3

MVCGGFACSKNCLCALNLLYTLVSLLLIGIAAWGIGFGLISSLRVVGVVIAVGIFLFLIALV
GLIGAVKHHQVLLFFYMIILLLVFIQFSVSCACLAQNQEQQQLLEVGNNTASARNDIQR
NLNCCGFRSVNPNDTCLASCVKSDHSCSPCAPIIGEYAGEVLRFVGGIGLFFSFTIELGVWL
TYRYRNQKDPRANPSAFL

Signal Peptide:

amino acids 1-34

Transmembrane domains:

amino acids 47-63, 72-95 and 162-182

FIGURE 16

TGATTGGAGCTGTAAAAAANTCTTCAGGTGTTGNATTTTTATATGATTATTCTGTAANT
TGTATTTATTGTTCAGTTNTGTATCTTGCCTTGCNTGAACCAGGAGCAACAGG
GTCAGNTNTGGAGGTTGGAAACAATACGGCAAGTGCTCGAAATGACATCCAGAGAAAT
NTAAACTGCTGTGGGTTCCGAAGTGTAAACCCAAATGACACCTGTNTGGCTAGCTGTGTTAA
AAGTGACCACTNGTGCCTGCCATGTGCTCCAATCATAGGAGAATATGCTGGAGAGGTTTGA
GATTGTTGGTGGCATTGGCCTGTTNTTCAGTTTACAGAGATCCTGGGTGTTGGCTGACC
TACAGATAACAGGAACCAAG

FIGURE 17

AATCCCAAATTCCCCAATTTTTGGNCTTTAGGGAAAGATGTGTTGTGGTAAAAAGTGT
TAGTATAAAAATGATAATTACTTGTAGTCTTTATGATTACACCAATGTATTCTAGAATAG
TTATGTCTTAGGAAATTGTGGTTAATTTGACTTTACAGGTAAGTGCAAAGGAGAAGTG
GTTTCATGAAATGTTCTAATGTATAATAACATTACCTTCAGCCTCCATCAGAATGGAACG
AGTTTGAGTAATCCAGGAAGTATATCTATGATCTTGATATTGTTTATATAATTGAAG
TCTAAAAGACTGCATTTTAAACAAGTTAGTATTAATGCCTGGCCACGTAGCAAAAGAT
ATTGATTATCTTAAAATTGTTAAATACCGTTTCATGAAAGTTCTCAGTATTGTAACAGC
AACTTGTCAAACCTAACGATATTGAATATGATCTCCATAATTTGAAATTGAAATCGTATT
GTGTGGAGGAAATGGCAATCTTATGTGTGCTGAAGGACACAGTAAGAGCACCAAGTTGTGCC
CCACTTGC

FIGURE 18

ATGATTATTCTGTTACTTGTATTTATTGTTCAGTTTATGGTATCTTGCCTGTTAGCCC
CTGAAACCAGGAGCAACAGGGNNCAGCTCCTGGAGGTTGGTGGCAACAATCACGGCCAAG
TGACTCCGCAAATGACATCCCAGAGAAACTCTAAACTGCTGTGGTTCCGAAGTGTAAACCC
AAATGACACCTGTCTGGCTNGCTGTGTTAAAGTGACCACTCGTGCTGCCATGTGCTCCAA
TCATAGGAGAATATGC

FIGURE 19

CAGTCACC**ATG**AAGCTGGCTGTGCCTCATGGCCTGGGCCCTACCTTCCCTGGTGTG
CTCTGGGTGGCCCAGATGCTACTGGCTGCCAGTTGAGACGCTGCAGTGTGAGGGACCTGT
CTGCACTGAGGAGAGCAGCTGCCACACGGAGGATGACTGACTGATGCAAGGGAAAGCTGGCT
TCCAGGTCAAGGCCTACACTTCAGTGAACCCCTCCACCTGATTGTGCTATGACTGGCTG
ATCCTCCAAGGTCCAGCAAGCCAGTTGAAGGGGACCTGCTGGTCTGCCTGCCAGGC
CTGGCAAGACTGGCCACTGACTCAGGTGACCTCTACCGAGATGGCTCAGCTCTGGTCCCC
CCGGGCCTAACAGGAATTCTCCATCACCGTGGTACAAAAGGCAGACAGCGGGCACTACCAC
TGCAGTGGCATCTCCAGAGCCCTGGTCTGGATCCCAGAAACAGCATCTGTTGTGGCTAT
CACAGTCCAAGAACTGTTCCAGCGCCAATTCTCAGAGCTGTACCCCTCAGCTGAACCCCAAG
CAGGAAGCCCCATGACCTGAGTTGTCAGACAAAGTTGCCCTGCAGAGGTAGCTGCCCGC
CTCCTCTTCCTCTACAAGGATGGAAGGATAGTGCAGAAAGCAGGGGCTCTCCTCAGAATT
CCAGATCCCCACAGCTTCAGAAGATCACTCCGGTCATACTGGTGTGAGGCAGCCACTGAGG
ACAACCAAGTTGAAACAGAGCCCCAGCTAGAGATCAGAGTGCAGGGTGTCCAGCTCT
GCTGCACCTCCCACATTGAATCCAGCTCCTCAGAAATCAGCTGCTCCAGGAACGTGCTCCTGA
GGAGGCCCTGGGCTCTGCCTCCGCCAACCCATCTCTGAGGATCCAGGCTTTCTT
CTCCTCTGGGATGCCAGATCCTCATCTGTATCACCAGATGGCCTTCTCAAACACATG
CAGGATGTGAGAGTCCTCGGTACCTGCTCATGGAGTTGAGGAATTATCTGCCACCA
GAAGCCTGGGACCACAAAGGCTACTGCTGA**ATAGA**AGTAAACAGTTCATCCATGATCTCACT
TAACCACCCAATAAAATCTGATTCTTATTTCTCTGCTGCACATATGCATAAGTA
CTTTACAAGTTGCCCAGTGTGTTAGAATAATGTTAGGTGAGTGTAAATAATT
ATATAAAGTGAGAATTAGAGTTAGCTATAATTGTGTATTCTCTTAAACACAAGAATT
TGCTGTCTAGATCAGGAATTCTATCTGTATATCGACCAGAATGTTGATTAAAGAGAA
CTAATGGAAGTGGATTGAATACAGCAGTCTCAACTGGGGCAATTGGCCCCCAGAGGACA
TTGGCAATTTGGAGACATTGGTCATTATACTTGGGGTTGGGGATGGTGGATGT
GTGTCTACTGGCATCCAGTAAATAGAAGCCAGGGTGCCTAAACATCCTATAATGCACAG
GGCAGTACCCACAAAGAAAAATACTGGCCAAATGTCAGTTGACTGAGTTGAGAAA
CCCCAGCCTAATGAAACCTAGGTGTTGGCTCTGGAAATGGACTTTGTCCCTCTAATT
TATCTCTTCCAGCCTCATCAGCTATTCTACTGACATACCACTTGTGCTATG
GTCCTTCTTAGTTCTAGTTGTATCCCTCAAAAGCCATTATGTTGAAATCCTAATCCCC
AAGGTGATGGCATTAAAGAAGTGGCCTTGGGAAGTGTAGATCAGGAGTGCAGAGCCCTC
ATGATTAGGATTAGTGCCTTATTAAAAAGGCCAGAGAGCTAACCTCCACCAT
ATGAGGACGTGGCAAGAAGATGACATGTATGAGAACCAAAACAGCTGTGCCAAACACCG
ACTCTGTCGTTGCCATTGATCTGAACCTCCAGCCTCCAGAACTATGAGAACAAATTCTGG
TTGTTGTAGCCTAA

FIGURE 20

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA40594
><subunit 1 of 1, 359 aa, 1 stop
><MW: 38899, pI: 5.21, NX(S/T): 0
MKLCVLMWALYLSLGVLWVAQMLLAASFETLQCEGPVCTEESSCHTEDDLTDAREAGFQV
KAYTFSEPFHLIVSYDWLILQGPAKPVFEGDLLVLRCQAWQDWPLTQVTFYRDGSALGPPGP
NREFSITVVQKADSGHYHCSGIFQSPGPGIPETASVVAITVQELFPAPILRAVPSAEPQAGS
PMTLSCQTKLPLRSAARLLFSFYKDGRIVQSRGLSSEFQIPTASEDHGSYWCEAATEDNQ
VWKQSPQLEIRVQGASSSAAPPTLNPAQKSAAPGTAPEEAPGPLPPPPTPSSEDPGFSSPL
GMPDPHLYHQMGLLLKHMQDVRVLLGHLLMELRELSGHQKPGTTKATAE
```

Signal sequence:

amino acids 1-17

Leucine zipper pattern sequence:

amino acids 12-33

Protein kinase C phosphorylation site:

amino acids 353-355

FIGURE 21

CCCACGCGTCCGCCAACGCGTCCGCCAACGGGTCCGCCAACCGCGTCCGGGCCACCAGAAGTT
TGAGCCTTTGGTAGCAGGAGGCTGGAAGAAAGGACAGAAGTAGCTCTGGCTGT**ATGGGG**
ATCTTACTGGCCTGCTACTCCTGGGCACCTAACAGTGGACACTTATGCCGTCCATCCT
GGAAGTGCCAGAGAGTGTAAACAGGACCTTGGAAAGGGGATGTGAATCTCCCTGCACCTATG
ACCCCTGCAAGGCTACACCCAAGTCTTGGTAAGTGGCTGGTACAACGTGGCTCAGACCCT
GTCACCATCTTCTACGTGACTCTTCTGGAGACCATATCCAGCAGGCAAAGTACCAAGGGCCG
CCTGCATGTGAGCCACAAGGTTCCAGGAGATGTATCCCTCCAATTGAGCACCCTGGAGATGG
ATGACCGGAGCCACTACACGTGTGAAGTCACCTGGCAGACTCCTGATGGCAACCAAGTCGTG
AGAGATAAGATTACTGAGCTCCGTGTCAGAAAACTCTCTGTCCTCAAGCCCACAGTGACAAC
TGGCAGCGGTTATGGCTTACGGTCCCCAGGGAATGAGGATTAGCCTTCAATGCCAGGCTC
GGGGTTCTCCTCCATCAGTTATTTGGTATAAGCAACAGACTAATAACCAGGAACCCATC
AAAGTAGCAACCCTAACGTACCTTACTCTCAAGCCTGCGGTGATAGCCGACTCAGGCTCTA
TTTCTGCACTGCCAACGGCCAGGTTGGCTCTGAGCAGCACAGCAGCATTGTGAAGTTGTGG
TCAAAGACTCCTCAAAGCTACTCAAGACCAAGACTGAGGCACCTACAACCATGACATACCC
TTGAAAGCAACATCTACAGTGAAGCAGTCCTGGGACTGGACCACTGACATGGATGGCTACCT
TGGAGAGACCAAGTGTGGCTGGGAGGAAAGAGCCTGCCTGTCTTGCCATCATCCTCATCATCT
CCTTGTGCTGTATGGTGGTTTACCATGGCTATATCATGCTCTGTCGGAAGACATCCAA
CAAGAGCATGTCTACGAAGCAGCCAGG**TAAGAAAGTCTCCTCTTCATTTGACCCGT**
CCCTGCCCTCAATTGATTACTGGCAGGAAATGTGGAGGAAGGGGGTGTGGCACAGACCC
AATCCTAACGGCGGAGGCCTTCAGGGTCAGGACATAGCTGCCCTCCCTCTCAGGCACCTT
CTGAGGTTTTGGCCCTCTGAACACAAAGATAATTAGATCCATCTGCCCTCTGCTTCC
AGAATCCCTGGGTGGTAGGATCCTGATAATTAAATTGGCAAGAATTGAGGCAGAAGGGTGGGA
AACCAGGACACAGCCCCAAGTCCCTCTTATGGGTGGTGGCTTGGGCATAGGGCACA
TGCCAGAGAGGCCAACGACTCTGGAGAAACCATGAGGGTGGCCATCTCGCAAGTGGCTGCT
CCAGTGATGAGCCAATTCCCAGAACTCTGGCAACAACTACTCTGATGAGCCCTGCATAGGA
CAGGAGTACAGATCATGCCAGATCAATGGCAACTACGCCGCTGTCGGACACAGTTCC
TCTGGATTATGAGTTCTGCCACTGAGGGAAAAGTGTCTGTTAAAATGCCCTTAGGC
CAGGATCTGCTGACATAATTGCCATTGCTAGTCAGTCCTGCCATGGCCTTCTGCT
ACCTCTCTCCTGGATAGCCAAAGTGTCCGCCTACCAACACTGGAGCCGCTGGAGTC
GGCTTGCCCTGGAATTGCCAGATGCATCTCAAGTAAGCCAGCTGCTGGATTGGCTCTGG
GCCCTCTAGTATCTGCCGGGGCTCTGGTACTCCTCTAAATACCAGAGGGAAAGATG
CCCATAGCACTAGGACTTGGTCATCATGCCACAGACACTATTCAACTTGGCATCTGCCA
CCAGAAGACCCGAGGGAGGCTCAGCTGCCAGCTCAGAGGACAGCTATATCAGGATCAT
TTCTCTTCTTCAGGGCCAGACAGCTTTAATTGAAATTGTTATTCACAGGCCAGGGTCA
GTTCTGCTCCTCCACTATAAGTCTAATGTTCTGACTCTCCTGGTGTCAATAATATCTA
ATCATAACAGC

FIGURE 22

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45416
><subunit 1 of 1, 321 aa, 1 stop
><MW: 35544, pI: 8.51, NX(S/T): 0
MGILLGLLLGHLTVDTYGRPILEVPESVTGPWKGDVNLPCTYDPLQGYTQVLVKWLVQRGS
DPVTIFLRDSSGDHIQQAKYQGRLHVSHKVPGDVSLQLSTLEMDDRSHYTCEVTWQTPDGNQ
VVRDKITELRVQKLSVSKPTVTTGSGYGFVQPQGMRISLQCQARGSPPISYIWYKQQTNNQE
PIKVATLSTLLFKPAVIADGSYFCTAKGQVGSEQHSDIVKFVVKDSSKLLKTKEAPTTMT
YPLKATSTVKQSWDWTTDMGYLGETSAGPGKSLPVFAIILIIISLCCMVVFTMAYIMLCRKT
SQQEHVYEAAR
```

Signal Sequence:

amino acids 1-19

Glycosaminoglycan attachment site:

amino acids 149-152

Transmembrane domain:

amino acids 282-300

FIGURE 23

GCGCCGGGAGCCCATCTGCCCCAGGGCACGGGGCGCGGGCGCTCCGCCGGCACAT
GGCTGCAGCCACCTCGCGCGACCCCGAGGGCGCGCCAGCTCGCCCGAGGTCCGTGG
GGCGCCCGGCCGCCGGAGCCAAGCAGCAACTGAGCGGGGAAGCGCCCGTCCGGGATC
GGGATGTCCCTCCTCCTCTGCTAGTTCTACTATGTTGAACCTGGGACTCA
CACTGAGATCAAGAGAGTGGCAGAGGAAAGGTCACTTGCCCTGCCACCATCAACTGGGC
TTCCAGAAAAAGACACTCTGGATATTGAATGGCTGCTCACCGATAATGAAGGGAACAAAAA
GTGGTGATCACTTACTCCAGTCGTATGTCTACAATAACTTGACTGAGGAACAGAAGGGCG
AGTGGCCTTGCTTCAATTCTGGCAGGAGATGCCTCTTGCGAGATTGAACCTCTGAAGC
CCAGTGATGAGGGCCGGTACACCTGTAAGGTTAAGAATTCAAGGGCGTACGTGTGGAGCCAT
GTCATCTAAAAGTCTTAGTGAGACCATCCAAGCCAAAGTGTGAGTTGAAGGAGAGCTGAC
AGAAGGAAGTGACCTGACTTGCAGTGTGAGTCATCCTCTGGCACAGAGCCCATTGTGTATT
ACTGGCAGCGAATCCGAGAGAAAGAGGGAGAGGATGAACGTCTGCCTCCAAATCTAGGATT
GACTACAACCACCCCTGGACGAGTTCTGCTGCGAGAATCTTACCATGTCTACTCTGGACTGTA
CCAGTGACAGCAGGCAACGAAGCTGGGAAGGAAGCTGTGGTGCGAGTAACTGTACAGT
ATGTACAAAGCATGGCATGGTGAGGAGCATGACAGGCATAGTGGCTGGAGCCCTGCTG
ATTTCTCTGGGTGCTGCTAATCCGAAGGAAGACAAAGAAAGATATGAGGAAGAAGA
GAGACCTAATGAAATTGAGAAGATGCTGAAGCTCCAAAGCCCGTCTGTGAAACCCAGCT
CCTCTCCTCAGGCTCTGGAGCTCACGCTCTGGTCTTCCACTCGCTCCACAGCAAAT
AGTGCCTCACGCAGCCAGGGACACTGTCAACTGACGCAGCACCCAGCCAGGGCTGGCCAC
CCAGGCATACAGCCTAGTGGGCCAGAGGTGAGAGGTTCTGAACCAAAGAAAGTCCACCATG
CTAATCTGACCAAAGCAGAAACCACACCCAGCATGATCCCCAGCCAGAGCAGAGCCTTCAA
ACGGTCT**TGA**ATTACAATGGACTTGAACCCCAGCTTCTAGGAGTCAGGTCTTGACTC
TTCTCGTCAATTGGAGCTCAAGTCACCAGCCACACAACCAGATGAGAGGTCACTAAGTAGCA
GTGAGCATTGCACCGAACAGATTAGCATGAGCATTCTCTTACAAATACCAAACAAGCAA
AGGATGTAAGCTGATTCACTGTAAAAGGCATCTTATTGTGCCTTAGCCAGAGTAAGGG
AAAGCAGGAGTCCAAATCTATTGTTGACCAGGACCTGTGGTGAGAAGGTTGGGAAAGGTG
AGGTGAATATACTAAAATTAAATGTGGATATTGTATCAGTGCTTGATTACAATT
TTCAAGAGGAATGGATGCTGTTGTAATTCTATGCATTCTGCAAACATTATTGGATT
ATTAGTTATTCAGACAGTCAGCAGAACCCACAGCCTTATTACACCTGTCTACACCATGTAC
TGAGCTAACCACTCTAACAAACTCCAAAAAGGAACATGTGTCTTCTATTCTGACTAAC
TTCATTGTCATAAGGTTGGATATTAAATTCAAGGGAGTTGAAATAGTGGAGATGGAGA
AGAGTGAATGAGTTCTCCACTCTATAACTATCTCACTATTGTATTGAGCCAAATAAC
TATGAAAGGAGACAAAATTGTGACAAAGGATTGTGAAGAGCTTCCATCTCATGATGTT
ATGAGGATTGTTGACAAACATTAGAAATATAATGGAGCAATTGTGGATTCCCTCAAAT
CAGATGCCTCTAACGGACTTCTGCTAGATATTCTGGAAGGAGAAATACAACATGTCATT
TATCAACGTCCTTAGAAAGAATTCTCTAGAGAAAAAGGGATCTAGGAATGCTGAAAGATTA
CCCAACATACCATTATAGTCTCTTCTGAGAAAATGTGAAACCAAGAATTGCAAGACTGG
GTGGACTAGAAAGGGAGATTAGATCAGTTCTTAAATATGTCAAGGAAGGTAGCCGGCA
TGGTGCAGGCACCTGTAGGAAAATCCAGCAGGTGGAGGTTGCAGTGAGCCGAGATTATGCC
ATTGCACTCCAGCCTGGGTGACAGAGCGGGACTCCGTCTC

FIGURE 24

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45419
><subunit 1 of 1, 373 aa, 1 stop
><MW: 41281, pI: 8.33, NX(S/T): 3
MSLLLLLVSYYVGTGTHTEIKRVAEEKVTLPCHHQLGLPEKDTLDIEWLTDNEGNQKV
VITYSSRHVYNNLTEEQKGRVAFASNFLAGDASLQIEPLKPSDEGRYCKVKNSGRYVWSHV
ILKVLVRPSKPKCELEGELTEGSDLTLQCESSSGTEPIVYYWQRIREKEGEDERLPPKSRID
YNHPGRVLLQNLTMYSGLYQCTAGNEAGKESCVVRTVQYVQSIGMVAGAVTGIVAGALLI
FLLVWLLIRRKDKERYEEEERPNEIREDAEAPKARLVKPSSSSGSRSSRSGSSSTRSTANS
ASRSQRTLSTDAAPQPGLATQAYSLVGPEVRGSEPKVHHANLTKAETTPSMIPSQSRAFQTV
```

Signal sequence:

amino acids 1-16

Transmembrane domain:

amino acids 232-251

FIGURE 25

GTCTCTCTCTCTCGCGCCAGTCCTCCCTGGTCTCCTCAGCCGTGCGGAGGAGACCCGGAGACGGGCTGAGTCGGCTCTCCCCGGCTGGGGCCTCGCCGCTGGGAGGTGCTGAGCGCCCTAGAGCCTCCCTGCCGCTCCCTCTGCCGGCGCAGCAGTCACATGGGTGTTGGAGGTAGATGGCTCCCGGCCGGAGGCGGGTGGATGCCGCTGGCAGAAGCAGCGCCATTCCAGCTGCCCGCGCAGGGGGCCGGCGCCCTGGCAGCTGCCATGGCCATGGCCAGGTGCTAACCTGTGACAAGTGTCCAGCAGGAACCTATGTCCTGAGCATTGTACCAACACAAGCTGCGCTTGCGAGCTTGCCTGTGGGGACCTTACCAAGGAGTGGATAGAGAAATGCCATGACTGTAGTCAGCCATGCCATGGCCAATGATTGAGAAATTACCTTGTGCTGCCCTGACTGACCGAGAAATGCACTGCCCACCTGGCATGTTCCAGTCTAACGCTACCTGTGCCCTACAGGAGTGTGCTGCCCTACAGGAGTGTGCTGCCCTAGGAGAAAGGGACAGAGACTGAGGATGTGCGGTGTAAGCAGTGTGCTGCCCTACAGATGTGCTGCCCTAGTGTGATGAAATGCAAAGCATAACAGACTGTCTGAGTCAGAACCTGGTGGTGAAGCCGGGACCAAGGGAGACAGACAACGCTGTGGCACACTCCCGTCCTTCGCCAGCTCCACCTCACCTCCCTGGCACAGCCATCTTCCACGCCCTGAGCACATGGAAACCCATGAAAGTCCCTTCCTCCACTATGTTCCCAAAGGCATGAACTCAACAGAACTCCAACTCTCTGCCCTCTGTTAGACCAAAGGACTGAGTAGCATCCAGGAAGGGACAGTCCCTGAAACACAAAGCTCAGCAAGGGGAGAAGACGTGAAACAGACCCCTCCAAACCTTCAGGAGTAGTCACCAACCCAGCAAGGCCCAACAGACACATCTGAAGGCTGTCGGTCCATGGAGGCCACTGGGGCGAGAAGTCAGCACGCCCATCAAGGGCCCAAGAGGGGACATCCTAGACAGAACCTACAAAGCATTGACATCAATGAGCATTGCTGGATGATTGTCCTTCTGCTGCTGGTCTTGTGGTATTGTGGTGTGAGTGCAGTATCCGGAAAAGCTGAGGACTCTGAAAAAGGGCCCGGCCAGGATCCAGTGCCTATTGTGGAAAAGGCAAGGCTGAAGAAATCCATGACTCCAACCCAGAACCCGGAGAAATGGATCTACTACTGCAATTGCCATGGTATCGATATCCTGAAGCTTGTAGCAGCCAAAGTGGAGCAGTGGAAAGATATCTATCAGTTCTTGCATGCCAGTGGAGAGGGAGGTTGCTGCTTTCTCAATGGTACACAGCCGACCACGAGCGGGCTACGCTCTGCAGCACTGGACCATCCGGGCCCCGAGGCCAGCCTCGCCAGCTAAATTGCGCCCTGCCAGCACGGAGAAACGATGTTGTGGAGAAGATTGCTGGGTGATGGAAGACACCAACCCAGCTGGAAACTGACAAACTAGCTCTCCGATGAGCCCCAGCCGCTTAGGCCGAGCCCCATCCCCAGCCCCAACGCGAAACTTGAGAATTGCCCTGACGGTGGAGCCTTCCCCCAGGACAAGAACAGGGCTTCTCGGGATGAGTCGGAGGCCCTTCTCGCTGTGACTCTACATCCAGCGGCTCTCCGCGTGGAGCAGCAGTGTGCTGCCCTTATTACCAAAGAAAAGAAGGACACAGTGTGCGGCAGGTACGCCCTGGACCCCTGTGACTTGAGCCTATCTTGATGACATGCTCCACTTCTAAATCCTGAGGAGCTGCCAGTGGGTGATTGAAGAGATTCCCCAGGCTGAGGACAAACTAGACCGGCTATTGAAATTATTGGAGTCAGAGCCAGGAAGCCAGGCCAGACCCTCCTGGACTCTGTTATAGCCATCTTCTGACCTGCTGTTAGAACATAGGGATACTGCATTGAAATTACTCAATTAGTGGCAGGGTGGTTTTAATTCTCTGTTCTGATTTTGTGTTGGGGTGTGTGTTGTGTTAACAGAGAATATGCCAGTGTGAGTTCTTCTCCCTCTCTCTCTCTCTTTTTAAATAACTCTTCTGGAGTTGTTATAAGCCTTGCCAGGTGTAAGTGTGAAATACCCACCACTAAAGTTTAAGTCCATATTCTCCTTGCCTCTTATGTTCAAGGATTATTCTGTCACCTTAAATTACTAACCTTACCATAAATGCACTGTCAGTGTGACTTTCCCACACACTGGATTGTGAGGCTCTAACCTCTTAAAGTATAATGGCATCTTGTAATCCTATAAGCAGTCCTTATGTCCTTAACATTACACACCTACTTTTAAAAACAAATATTACTATTCTTATTGTTGCTCTTATAAAATTCTTAAAGATTAAGAAAATTAAAGACCCCATTGAGTTACTGTAATGCAATTCAACTTGGAGTTATCTTAAATATGTCCTGTTGATAGTCATTCTGCTGAGGACTCTGAGGACACACTTGGACACTTGGCTGAGGACACTTGGCTGAGGAAATGCAAGGCTTCAAGGAGCAGGATTTGCTATTGTCATTAAAGTGTGTTGCTCTTCTGAAACTTGGCCACCAAAAGAAGTGAACATTCTGAGGTTAGGATTTGAGCTGTCAGGCTGAGGAGCTTGGAGGATCTTAAAGATGTCTGTTTTTTATATACACATAATCAATAGGCTCAACTGTAAGAACCTTGTGCTGATATAATTGCAACTATGCTCCCTTGTGAGGGTGGGGTTGTGGTAGTGGTGAAGGACCGATATCAGAAAAAATGCCTCAAGTGTACTAATTAAACATTAGGTGTTGTTAAAAAA

FIGURE 26

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52594
><subunit 1 of 1, 655 aa, 1 stop
><MW: 71845, pI: 8.22, NX(S/T): 8
MGTPSSSTALASCSRIARRATATMIAGSLLLLGFLSTTTAQPEQKASNLIGTYRHVDRATG
QVLTCDKCPAGTYVSEHCTNTSLRVCSSCPVGTFRHENGIEKCHDCSQPCPWPMIEKLPCA
ALTDRECTCPGFMQSNATCAPHTVCAPVGWVKKGTETEDVRCKQCARGTFSDVPSSVMKC
KAYTDCLSQNLVVIKPGTKETDNVCGLPSFSSSTSPSPGTIAIFPRPEHMETHEVPSSTYVP
KGMNSTESNSSASVRPKVLSSIQEGLVDPNTSSARGKEDVNKTLPNLQVNHQQGPHHRHIL
KLLPSMEATGGEKSSTPIKGPKRGHPRQNLHKHFDINEHLPWMIVLFLLLVLVVIIVVCSIRK
SSRTLKKGPRQDPSAIVEKAGLKKSMPTQNREKWIYYCNGHGIDILKLVAAQVGSQWKDIY
QFLCNASEREVAAFSNGYTADHERAYAALQHWTIRGPEASLAQLISALRQHRRNDVVEKIRG
LMEDTTQLETDKLALPMSPSPLSPSPPIPSNAKLENSALLTVEPSPQDKNKGFFVDESEPLL
RCDSTSSGSSALSRNGSFITKEKKDTVLRQVRLDPCLQPIFDDMLHFLNPEELRVIEEIPQ
AEDKLDRLFEIIGVKSQEASQTLDSVYSHLPDLL
```

Signal sequence:

amino acids 1-41

Transmembrane domain:

amino acids 350-370

FIGURE 27

ATGGGAAGCCAGTAACACTGTGGCCTACTATCTCTTCCGTGGTGCCATCTACATTTTGGGA
CTCGGGAAATTATGAGGTAGAGGTGGAGGCAGCCGGATGTCAGAGGTCTGAAATAGTCAC
CATGGGGAAAATGATCCGCCTGCTGTTGAAGCCCCCTCTCATTCCGATCGCTTTGGCC
TTGATGATTGAAAATAAGCCTGTTGCACCAGATGCAGATGCTGTTGCTGCACAGATCCTG
TCACTGCTGCCATTGAAGTTTTCCAATCATCGTATTGGATCATGCATTGATATTAGC
ACTGGCCATTGGCTGGGCATCCACTTCGACTGCTCAGGGAAAGTACAGATGTCGCTCATCCT
TTAAGTGTATCGAGCTGATAGCTCGATGTGACGGAGTCTGGATTGCAAAGACGGGAGGGAC
GAGTACCGCTGTGTCGGGTGGTCAGAAATGCCGTCTCAGGTGTTCACAGCTGCTTC
GTGGAAGACCATGTGCTCCGATGACTGGAAGGGTCACTACGCAAATGTTGCCTGTGCCAAC
TGGGTTTCCAAGCTATGTGAGTTGAGATAACCTCAGAGTGGACTCGCTGGAGGGCAGTTC
CGGGAGGAGTTGTGTCATCGATCACCTTGCAGATGACAAGGTGACTGCATTACACCA
CTCAGTATATGTGAGGGAGGGATGTGCTCTGCCACGTGGTTACCTGCAGTGCACAGCCT
GTGGTCATAGAAGGGCTACAGCTCACGCATCGTGGGTGAAACATGTCCTGCTCTCGCAG
TGGCCCTGGCAGGCCAGCCTCAGTTCCAGGGTACCCACCTGTGCGGGGCTCTGTCATCAC
GCCCTGTGGATCATCACTGCTGCACACTGTGTTATGACTTGTACCTCCCCAAGTCATGGA
CCATCCAGGTGGGTCTAGTTCCCTGTTGACAATCCAGCCCCATCCACTTGGTGGAGAAG
ATTGTCTACCACAGCAAGTACAAGCCAAAGAGGCTGGCAATGACATCGCCTTATGAAGCT
GGCCGGGCCACTCACGTTCAATGAAATGATCCAGCCTGTGCGCTGCCAACTCTGAAGAGA
ACTTCCCCGATGGAAAAGTGTGCTGGACGTCAAGGATGGGGGCCACAGAGGATGGAGGTGAC
GCCTCCCCGTCTGAACCACGCGGCCGTCCCTTGATTCCAACAAGATCTGCAACCACAG
GGACGTGTACGGTGGCATCATCTCCCCCTCATGCTCTGCGCGGGCTACCTGACGGTGGCG
TGGACAGCTGCCAGGGGACAGCGGGGGCCCTGGTGTCAAGAGAGGAGGCTGTGGAAG
TTAGTGGGAGCGACCAGCTTGGCATGGCTGCGAGAGGTGAACAAGCCTGGGTGTACAC
CCGTGTACCTCCTGGACTGGATCCACGAGCAGATGGAGAGAGACCTAAAAACCT**GA**
GAGGAAGGGGACAAGTAGCCACCTGAGTCTGAGGTGATGAAGACAGCCGATCCTCCCCT
GGACTCCCGTGTAGGAACCTGCACACGAGCAGACACCCCTGGAGCTGAGTTCCGGCACCA
GTAGCAGGCCGAAAGAGGCACCCCTCATGATTCCAGCACAACCTCAAGCTGCTTTTT
GTTTTTGTGTTTGAGGTGGAGTCTCGCTCTGTTGCCAGGCTGGAGTGCAGTGGCGAAA
TCCCTGCTCACTGCAGCCTCCGCTTCCCTGTTCAAGCGATTCTCTGCCTCAGCTCCCCA
GTAGCTGGGACCACAGGTGCCGCCACACACCCAAACTAATTTGTATTTAGAGAGAC
AGGGTTTACCATGTTGCCAGGCTGCTCTCAAACCCCTGACCTCAAATGATGTCCTGCTT
CAGCCTCCCACAGTGTGGATTACAGGCATGGCCACCACGCCCTAGCCTCACGCTCCTTTC
TGATCTTCACTAAGAACAAAAGAAGCAGCAACTTGCAAGGGGGCCTTCCCCTGGTCCAT
CTGGTTTCTCTCCAGGGTCTGAAAATTCTGACGGAGATAAGCAGTTATGTGACCTCAG
TGCAAAGCCACCAACAGCCACTCAGAAAAGACGCACCAGCCCAGAAGTGCAGAACTGCA
ACTGCACGTTTCACTCTAGGGACCAGAACCAAACCCACCCCTTCTACTTCCAAGACTTAT
TTTCACATGTGGGAGGTTAATCTAGGAATGACTGTTAAGGCCTATTTCATGATTCTT
TGTAGCATTTGGTGTGCTGACGTATTATTGTCCTTGATTCCAAATAATATGTTCCCTCCCT
CATTGTCTGGCGTGTGCGTGGACTGGTGACGTGAATCAAATCATCCACTGAAA

FIGURE 28

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45234

><subunit 1 of 1, 453 aa, 1 stop

><MW: 49334, pI: 6.32, NX(S/T): 1

MGENDPPAVEAPFSFRSLFGLDDLKISPVAPDADAVAAQILSLLPLKFFPIIVIGIIALILA
LAIGLGIHFDCSGKYRCRSSFKCIELIARCDGVSDCKDGEDEYRCVRVGGQNAVLQVFTAAS
WKTMCSDDWKGHYANVACAQLGFPSYVSSDNLRVSSLEGQFREFVSIDHLLPDDKVTALHH
SVYVREGCASGHVVTLQCTACGHRRGYSSRIVGGNMSLLSQWPWQASLQFQGYHLCGGSVIT
PLWIITAHCVYDLYLPKSWTIQVGLVSLLDNPAPSHLVEKIVYHSKYKPKRLGNDIALMKL
AGPLTFNEMIQPVCLPNSEENFPDGKVCWTSGWGATEDGGDASPVLNHAAPVLIISNKICNHR
DVYGGIISPSMLCAGYLTGGVDSCQGDSGGPLVCQERRLWKLVGATSGIGCAEVNKPGVYT
RVTSFLDWIHEQMERDLKT

Signal Peptide:

amino acids 1-20

Transmembrane domain:

amino acids 240-284

FIGURE 29

CCACCGCGTCCGTCTAGTCCCCGGCCAACCGGACAGTTGCTCATTATTGCAACGGTCAAGGCTGGCTTGT
GCCAGAACGGCGCGCGCGCGCACGCACACACACGGGGGGAAACTTTTAAAAATGAAAGGCTAGAAGA
GCTCAGCGGGCGCGCGGGCGCTGCGCAGGGCTCCGGAGGCTGACTCGCCGAGGAGGAAATCCCTCCGGTCGCGA
CGCCCGGCCCCGGCTCGCGCCCGCTGGGATGGTGCAGCGCTCGCCGCGGGCCCGAGAGGCTGACTGAG
GCCGGCGACGATGGCAGCGCGCCCGCTGCCGTGCCCCGCGCCCTCTGCTCGCCCTGGCGGTGCTCT
GCTCGCGCCCTGCGAGGGCGAGGGGTGAGCTTATGAACCAAGGAAGAGCTGATGAAGTTGTCAGTC
TCGGAGTGGGACCTCTGGATCCCAGTGAAGAGCTTCGACTCCAAGAATCATCCAGAAGTGTGAATATTGACT
ACAACGGGAAAGCAAAGAACTGATCATAAATCTGAAAGAAATGAAGGTCTATTGCCAGCAGTTCACGAAAC
CCACTATCTGCAAGACGGTACTGATGTCTCCCTCGCTCGAAATTACACGGGTCACTGTTACTACCATGGACATGT
ACGGGGATATTCTGATTCAAGCAGTCAGTCTCAGCACGTGTTCTGGTCTCAGGGACTTATTGTGTTGAAATGA
AAGCTATGCTTACAACATGAAACAGATACAAACTCTCCCAGCGAAGAAGCTGAAAGCGT
CCGGGATCATGTGGATCACATCACAAACACACCAAACTCGCTGCAAAGAATGTGTTCCACCACCCCTCAGAC
ATGGGCAAGAAGGCATAAAAGAGAGACCTCAAGGCAACTAAGTATGTGGAGCTGGTGATCGTGGCAGACAACCG
AGAGTTTCAGAGGCAAGGAAAAGATCTGGAAAAAGTTAAGCAGCGATTAAATAGAGATGCTAATCACGTTGACAA
GTTTACAGACCACTGAACATTGGATCGTGTGGTAGGGCTGGAAGTGTGGAATGACATGGACAAATGCTCTGT
AAGTCAGGACCCATTCAACAGCCATGAATTCTGGACTGGAGGAAGATGAAGCTTCTACCTCGCAAATCCCA
TGACAATGCGCAGCTTGTCACTGGGTTATTCCAAGGGACCATCGCATGGCCCAATCATGAGCATGTG
CACGGCAGACCGACTGGGAAATTGTCACTGGACCATTCAGACAAATCCCCTGGTGCAGCGTGACCCCTGGCACA
TGAGCTGGGCCACAATTGGGATGAATCATGACACACTGGACAGGGCTGTAGCTGTCATGGCGGTTGAGAA
AGGAGGCTGCATCATGAACGCTTCCACGGGTACCCATTCCATGGTGTCACTGGAGGACTT
GGAGACCAAGCTGGAGAAAGGAATGGGGTGTGCCTTTAACCTGCCGAAGTCAGGGAGTCTTCGGGGCCA
GAAGTGTGGAACAGATTGTGAAAGAAGGAGAGGAGGTGTGACTGTGGGAGCCAGAGGAATGTATGAATCGCTG
CTGCAATGCCACCACCTGTACCCCTGAAGCCGGACGCTGTGCGCACATGGCTGTGTCAGACTGCCAGCT
GAAGCCTGCAGGAACAGCGTGCAGGGACTCCAGCAACTCCTGTGACCTCCAGTGTCAATGGCAGCC
TCACTGCCAGCAATGTGTACCTGCACGATGGCACTCATGTCAGGATGTGGACGGCTACTGCTACAATGGCAT
CTGCCAGACTCACGAGCAGCAGTGTGTCACGCTCTGGGACAGGTGCTAACCTGCCCTGGATCTGCTTGA
GAGAGTCATTCTGCAGGTGATCCTTATGCAACTGTGGCAAAGTCTCGAAGAGTTCTTTGCCAAATGCCAGAT
GAGAGATGCTAAATGTGAAAATCCAGTGTCAAGGAGGTGCCAGCCGGCAGTCATTGGTACCAATGCCATT
CATAGAAACAAACATCCCTCTGCAGCAAGGAGGCCATTGTGCCGGGGACCCACGTGTACTTGGCGATGA
CATGCCGACCCAGGGTTGTGCTGCAGGCACAAAGTGTGCACTGGAAAAATCTGCCCTGAATCGTCAATGTCA
AAATATTAGTGTCTTGGGTTCACGAGTGTCAATGCACTGCCACGGCAGAGGGTGTGCAACAAACAGGAAGAA
CTGCCACTGCGAGGCCACTGGCACCTCCCTGTGACAAGTTGGTTGGAGGAAGCAGACAGCGGCC
CATCCGGCAAGCAGAACAGCAAGGAGGCCATTGTGCAAGACTCCACAGGGAGGCCAGGGAGCCCGTGG
ATCGCAGGAGCAGCGCTACTGCCCTACTGACACTCATCTGAGCCCTCCCATGACATGGAGACCGTGA
CTGCTGCAGAGGAGGTACCGCTCCCAAGGCCCTGTGACTGGCAGCATTGACTCTGTGGCTTGCCATCGTT
TCCATGACAACAGACACAACACAGTTCTGGGCTCAGGAGGGAAAGTCCAGCCTACAGGACAGTCTGCA
CACTGCAAGGAAGGGCAGCGACTCTGGTGTGAGCTTCTGCTAAACATGGACATGCTTCACTGCTC
AGAGTAGCAGGTTTACACTCTGGCAGGCCAGGCCCTGCAAGGAGGAAGGACTCAAAAGTCTGGCTTTC
ACTGAGCCTCCACAGCAGTGGGGAGAAGCAAGGGTTGGGCCAGTGTCCCCTTCCCCAGTGA
TGGCAGCCCTGATGACTGGTCTCTGGCTGCAACTTAATGCTGTGATATGGCTTTAGCATTATTATGAAA
AGCAGGGTTTGTGTTAAATTATCAGAGACCTGCCACCCATTCCATCTCCATCCAAGCAA
TGAAACAAACTGGAGAAGAAGGTAGGAGAAGGGCGGTGAACCTGGCTCTTGCTGTGACATGCGTGA
AGTACTCAGGTTTGAGGGTTGCAAGAACGCCACAGAGTCACCAACCCTCATTAA
TGTTAAAAGTGAAACAAATGTAAGAGCCTAACTCCATCCCCGTGCCATTACTGC
GATAAAATAGAGTGCATTT
GAAAT

FIGURE 30

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49624
><subunit 1 of 1, 735 aa, 1 stop
><MW: 80177, pI: 7.08, NX(S/T): 5
MAARPLPVSPARALLALAGALLAPCEARGVSLWNQGRADEVVSASVRSGDLWIPVKSFD SK
NHPEVLNIRLQRESKELIINLERNEGLIASSFTETHYLQDGTDVSLARNYTGHCYYHGHVRG
YSDSAVSLSTCSGLRGLIVFENESYVLEPMKSATNRYKLFPAKKLKSRGSCGSHHNTPNLA
AKNVFPPPSQTWARRHKRETLKATKYVELVIVADNREFQRQGKDL EKVVKQRLIEIANHVDKF
YRPLNIRIVLVGVEVWNDMDKCSVSQDPFTSLHEFLDWRKMKLLPRKSHDNAQLVSGVYFQG
TTIGMAPIMSMCTADQSGGIVMDHSDNPLGAAVTLAHELGHNFGMNHDTLRGCSCQMAVEK
GGCIMNASTGYPFFPMVFSSCSRKDLETSLEKGMGVCLFNLP EVRESFGGQKCGNRFVEEGEE
CDCGEPEECMNRC CNATTCTLKPDAVCAHGLC EDCQLKPAGTACRDSSNSCDLPEFCTGAS
PHCPANVYLHDGHSCQDV DGYC YNGICQTHEQQCVTLWGP GAKPAPGICFERVNSAGDPYGN
CGKVKSSFAKCEMRDAKCGKIQ CQGGASRPVIGTNAV SIETNIPLQQGGRILCRGTHVYLG
DDMPDPGLVLAGTKCADGKICLNRQCQNISVFGVHECAMQC HGRGV CNNRKNCHCEAHWAPP
FCDKFGFGGGSTD SGPIRQAEARQEAAESNRERGQGQEPVGSQE HASTASLT LI
```

Signal peptide:

amino acids 1-28

FIGURE 31

TCCCAAGGCTTCTTGGATGGCAGATGATTNTGGGTTTGCAATTGTTCCCTGACAACGAAA
ACAAAACAGTTGGGGTTCAAGGAGGGAAANTCCAGCCTACCCAGGAAGTTGCAGAAACA
GTGCAAGGAAGGGCAGGANTCCTGGTGAGNTTTGNTAAAACATGGACATGNTTCAGTG
CTGCTCNTGAGAGAGTAGCAGGTTACCACTTTGGCAGGCCAGCCCTGCAGCAAGGAGGA
AGAGGACTCAAAAGTTGGCCTTCACTGAGCCTCACAGCAGTGGGGAGAACAGCAAGGGTT
GGGCCAGTGTCCCCTTCCCCAGTGACACCTCAGCCTGGCAGCCCTGATAACTGGTNTNT
GGCTGCAANTTAATGCTNTGATATGGCTTTAGCATTATTATGAAAATAGCAGGGTTTT
AGTTTTAATTATCAGAGACCCTGCCACCCATTCCATNTCCATCCAAG

FIGURE 32

CATCCTGCAACATGGTAAACCACGCCCTGGCTAATTTGTTGTTAGAGATGGGA
TTTCACCGTGTAGCCAGGATTGTCTCAATCTGACCTCATGATCTGCCCGCCTGGCCTCCC
AAAGTGCTGGATTACAGGCGAGTGCAACCACACCCGGCCACAAACTTTAAGAAGTTAAT
GAAACCATACTTTACATTTAATGACAGGAAAATGCTCACAATAATTGTTAACCCAAAA
TTCTGGATACAAAAGTACAATCTTACTGTGTAATACATGTATATGACTATATGAAAATA
TACCAAATATCAATAACTTATCTCTGGTAAAACCTCTCATACCCGTGCTAACAA
CTTTAACAAAAAATTGCATCACTTTAAGAATCAAGAAAATTCGAAGGTCATATGGG
ACAGAAAAAAACCAAGGGAAAATCACGCCACTTGGGAAAAAGATTGAAATCTGCCT
TTTATAGATTGTAATTAATAAGGTCCAGGCTTCTAAGCAACTAAATGTTGTTGCA
AACAAAGTACTGTCTGGATGTAGGAGGAAAGGGAGTGATGTCAGTGCATTATGATGCC
TTGAATATAAGACCCACTTGCTATCTCCCTGCACCAGCCAGGAGCCACCCATCCTCCAGC
ACACTGAGCAGCAAGCTGGACACACGGCACACTGATCCAAATGGTAAGGGATGGTGGCGA
TGCTCATTCTGGTCTGCTACTTCTGGCGCTGCTCCTACCCGTGCAGGTTCTTCATTGTT
CCTTAACCAGTATGCCGGAAGCTACTGCAGCCAAACCACAAAGCCTCCAACAGTGC
ACAGCCTACAGCCGGTCTCCTGTGGCTTGCTTGCCTTCTACATCTTACCATTAAGAGG
CAGGTCAAGAACAGCTACAGTTCTCCAACCCATACACTAAAACCGAATCCAATGGTGC
AGAAGTTCAATGTGGCAAGGAAAAAACAGGTCTTCATCAAATCTACTAATTCACTC
ATTAACAGAGAACGCTTGAGAGTCTCAAACGGACTGGTTAAAGAGCATCTGAAGGATT
GACTAGATGATAATGCCTGTACTCCCAGTACTTGGGAGGCCTAGGCCGGGATCACCTG
AGGTCAAGGAGTTGAGACTAACCTGGCAAATGGTGAACCCATCTGTACTAAAATACA
AATATTGACTGGCGTGGTGGTGAGTGCCTGTGATCCCAGCTACTCAGGTGGCTGAAGCAGG
ACAATCACTGAACTCAGGAGGCAGAGGTTGCAGTGAGCTGAGATCGCGCTACTGCACTCTA
GCCCTAGCCTGGCAACAGAGTGAGACTTCGTCTCAAAAAAAAAAGCCAAGTGCAGTGGCT
CACGCCTGTAATCCCGCACTTGGGAGGCCAGGTGGCGGATCACGAGGTCAAGGAGATCA
AGACCATCCTGGCTAATACAGTGAAACCCCTGTCTACTAAAATACAACAAATTAGCC
GATGGTGGCAGGCACCTGGAGTCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATAGCGTGA
CTCAGGAGGCAGGAGCTGCACTGAGCCAGATTGCGCTACTGCACTCCAGCCTGGCGACAG
CGCGAGACTCCGTCTCAAAAAAAAAAAAAAA

FIGURE 33

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48309

><subunit 1 of 1, 67 aa, 1 stop

><MW: 6981, pI: 7.47, NX(S/T): 0

MGKGMVAMLILGLLLLALLLPVQVSSFVPLTSMPEATAAETTKPSNSALQPTAGLLVVLLAL
LHLYH

Signal peptide:

amino acids 15-27

FIGURE 34

GGCGCGCGAGAGCGCGCCCAGCCCCGCCGCGATGCCCGCGCCAGGACGCCCTCCCGCTGCTGGCCCGC
CGCGGCCCTGACTCGCCTGCTGCTGCTGGCCATGGCGCGGGCGCTGGGGGCCCGGGCCAGG
AGGCGGCCGGCGGCCGGCGGACGGGCCCGCGCAGACGGCGAGGACGGACAGGACCCGACAGCAAGCACC
TGTACACGGCCGACATGTTACGCACGGGATCCAGAGCGCCGCACCTCGTCATGTTCTCGCCCTGGTGTG
GACACTGCCAGCGCTGAGCCACTTGGAAATGACCTGGAGACAAATAACACAGCATGGAAGATGCCAAAGTCT
ATGTGGCTAAAGTGGACTGCACGCCACTCCGACGTGCTCCGCCAGGGGTGCGAGGATAACCCACCTAA
AGCTTTCAAGCCAGGCCAAGAAGCTGTGAAGTACCAAGGGTCTCGGGACTTCCAGACACTGGAAAATGGATGC
TGCAGACACTGAACGAGGAGCCAGTGACACCAGAGCCGAAGTGGAAACGCCAGTGCCCCGAGCTCAAGCAAG
GGCTGTATGAGCTCTCAGCAAGCAACTTGGAGCTGCACGTTGACAAGGGGACCACTTATCAAGTTCTCGCTC
CGTGGTGTGGTCACTGCAAAGCCCTGGCTCAACCTGGAGCAGCTGGCTCTGGCCCTGAACATTCCGAAACTG
TCAAGATTGCAAGGTTGATTGTACACAGCACTATGAACACTCTGCTCCGAAACCAAGGTTGCTATCCACTC
TTCTCTGGTCCGAGATGGAAAAAGGTGGATCAGTACAAGGAAAGCGGATTGGAGTCACTGAGGGAGTACG
TGGAGTCGAGCTGCAGCGCACAGAGACTGGAGCGACGGAGACCGTCACGCCCTCAGAGGCCCGGTGCTGGCAG
CTGAGCCCGAGGCTGACAAGGGCACTGTGTTGGCACTCACTGAAAATAACTTCGATGACACCATTGAGAAGGAA
TAACCTCATCAAGTTTATGCTCCATGGTGTGGTCATTGTAAGACTCTGGCTCTACTGGAGGAACCTCTCA
AAAAGGAATTCCCTGGTCTGGGGGGTCAAGATGCCGAAGTAGACTGCACTGCTGAACGGAAATATCTGAGCA
AGTATTGCGTACAGGGTACCCCCACGTTATGCTTTCCGAGGGGAAGAAAGTCAGTGAGCACAGTGGAGGCA
GAGACCTTGACTCGTTACACCGCTTGTGCTGAGCCAAGCGAAAGACAACTT~~TAGGA~~ACACAGTGGAGGTCAC
CTCTCTGCCAGCTCCGACCCCTGCGTTAGGAGTTCACTGCCCACAGAGGCCACTGGGTTCCAGTGGCT
GTTCAAGAAAGCAGAACATACTAACGGCTGAGGTATCTCTTGTGTTGTTCCAGGAAACACACTCTACAG
ATTCTTTATTAAAGTTAAGTTCTAAGTAATGTGTAACTCATGGTCACTGTGTAACATTTCACTGCGATA
TATCCCCTTGACCTTCTCTGATGAAATTACATGGTTCTTGAGACTAAATAGCGTTGAGGAAATGAAA
TTGCTGGACTATTGTTGGCTCCTGAGTTGAGTGATTGAAAGAACATCAAAGCATAGTTACCTGC
CCACGAGTTCTGAAAGGTGGCTTGTGGCAGTATTGACGTTCTGATCTTAAGGTACAGTTGACTCAATAC
TGTGTTGGTCCGTAGCATGGAGCAGATTGAAATGCAAAACCCACACCTCTGGAAGATACTTCACGGCCGCTGC
TGGAGCTTCTGTTGCTGTAATACTCTCTCAGTGTGAGAGGTTAGCGTGTGAAAGCAGCGTTACTCTGACC
GTGCCTGAGTAAGAGAATGCTGATGCCATAACTTATGTGTCATACTGTCAAATCAGTTACTGTTCAGGGAT
CCTCTGTTCTACGGGGTAAACATGTCTTAGTCCTCATGTTAACACGAAGCCAGAGGCCACATGAACTGT
TGGATGTCCTTAGAAAGGGTAGGCATGGAAAATTCCACGAGGCTCATCTCAGTATCTCATTAACCTCATTGA
AAGATTCCAGTTGATTGTCACCTGGGGTACAAGACCCAGACAGGCTTCCAGGCCTGGGTATCCAGGGAGGC
TCTGCAGCCCTGCTGAAGGGCCCTAACTAGAGTTCTAGAGTTCTGATTCTGTTCTCAGTAGTCCTTAGGAGG
CTTGCTATACTGGTCTGCTCAAGGAGGTGACCTCTAATGTATGAAAGATGGGATGCATTGATCTCAAGAC
CAAAGACAGATGTCAGTGGCTGCTGGCCCTGGTGTGCACGGCTGTGGCAGCTGTTGATGCCAGTGTCTCTA
ACTCATGTCCTTGTGATTAACACCTCTATCTCCCTGGGAAATAAGCACATACAGGCTTAAGCTCTAAAGATA
GATAGGTGTTGTCCTTTACCATCGAGCTACTTCCCTAATAACCAACTTGTGATCCAACACTCTCACCCACCT
CCCATACGCAAGGGATGTGGATACTGGCCAAAGTAACTGGTAGGAATCTTAGAAACAAGACCAACTTATA
CTGTCGTTGAGGAGAAGATAACAGCAGCATCTGACCGAGCCTCTGCCTTAAAGGAAATCTTATTAAATCAG
TATGGTTCACAGATAATTCTTTTTAAAAAAACCAACCTCTAGAGAAGCACAACGTCAAGAGTCTGTACA
CACAACTTCAGCTTGACATCACGAGTCTGTATTCCAAGAAAATCAAAGTGGTACAATTGTTGTTACACTAT
GATACTTTCTAAATAACTCTTTTTTAA

FIGURE 35

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA46776
><subunit 1 of 1, 432 aa, 1 stop
><MW: 47629, pI: 5.90, NX(S/T): 0
MPARPGRLLPLLLARPAALTALLLLGHGGGGRWGARAQEAAAAADGPPAADGEDGQDPHS
KHLYTADMFTHGIQSAAHFVMMFFAPWCGHCQRLQPTWNDLGDKYNSMEDAKVYVAKVDCTAH
SDVCSAQGVRYPTLKLFKPGQEAVKYQGPRDFQTLLENWMLQTLNEEPVTPEPEVEPPSAPE
LKQGLYELSAASFELHVAQGDHFIKFFAPWCGHCKALAPTWQLALGLEHSETVKIGKVDCT
QHYELCSGNQVRYPTLLWFRDGKKVDQYKGKRDLESLEYVESQLQRTETGATETVTPSEA
PVLAEEPEADKGTVLALTENNDDTIAEGITFIKFYAPWCGHCKTLAPTWEELSKKEFPGLA
GVKIAEVDCTAERNICSKYSVRGYPTLLLFRGGKKVSEHSGGRDLDLHRFVLSQAKDEL
```

Signal sequence:

amino acids 1-32

FIGURE 36

CTTTCTGAGGAACCACAGCAATGAATGGCTTGCATCCTGCTCGAAGAAACCAATTAT
CCTCCTGGTACTATTCCTTGCAAATTCAAGAGTCTGGGCTGGATATTGATAGCCGTCTA
CCGCTGAAGTCTGTGCCACACACACAATTCAACCAGGACCCAAAGGAGATGATGGTAAAAA
GGAGATCCAGGAGAAGAGGGAAAGCATGGCAAAGTGGACGCATGGGCCGAAAGGAATTAA
AGGAGAACTGGGTGATATGGGAGATCAGGGCAATATTGGCAAGACTGGGCCATTGGAAAGA
AGGGTGACAAAGGGAAAAGGTTGCTTGAATACCTGGAGAAAAGGCAAAGCAGGTACT
GTCTGTGATTGTGGAAGATACCGAAATTGTTGGACAATGGATATTAGTATTGCTCGGCT
CAAGACATCTATGAAGTTGTCAAGAATGTGATAGCAGGGATTAGGGAACTGAAGAGAAAT
TCTACTACATCGTCAGGAAGAGAACTACAGGGAACTCCCTAACCCACTGCAGGATTCGG
GGTGAATGCTAGCCATGCCAAGGATGAAGCTGCCAACACACTCATCGCTGACTATGTTGC
CAAGAGTGGCTTCTTCGGGTGTTCATGGCGTGAATGACCTTGAAAGGGAGGGACAGTACA
TGTCCACAGACAACACTCCACTGCAGAACTATAGCAACTGGAATGAGGGGGAACCCAGCGAC
CCCTATGGTCATGAGGACTGTGAGATGCTGAGCTCTGGCAGATGGAATGACACAGAGTG
CCATCTTACCATGTACTTGTCTGTGAGTTCATCAAGAAGAAAAGTAACTTCCCTCATCCT
ACGTATTTGCTATTCCTGTGACCGTCATTACAGTTATTGTTATCCATCCTTTTCCTG
ATTGTAACATTTGATCTGAGTCACATAGCTAGAAAATGCTAAACTGAGGTATGGAGCCT
CCATCATCAAAAAAAAAAAAAAA

FIGURE 37

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50980
><subunit 1 of 1, 277 aa, 1 stop
><MW: 30645, pI: 7.47, NX(S/T): 2
MNGFASLLRRNQFILLVLFLLQIQSLGLDIDSRPTAEVCATHTISPGPKGDDGEKGDPGEEG
KHGKVGRMGPKGIKGELGDMGDQGNIGKTGPIGKKGDKGEKGLLGIPGEKGKAGTVCDGRY
RKFVGQLDISIARLKTSMKFVKVNVIAAGIRETEEKFYIIVQEEKNYRESLTHCRIRGGMLAMP
KDEAANTLIADYVAKSGFFRVFIGVNDLEREGQYMSTDNTPLQNYSNWNEGEPSDPYGHEDC
VEMLSSGRWNDTECHLTMYFVCEFIKKKK
```

Signal peptide:

amino acids 1-25

FIGURE 38

GGTTCTATCGATTGAATTCGCCACACTGGCCGGATCCTCTAGAGATCCCTGACCTCGAC
CCACCGCGTCCGCTGCTCTCCGCCGTGGAGTGGTGGGGCTGGGTGGGAATGGCGTGT
GCCAGCGACGCGCGCTCCCTGGAAGGAGAAGTCTCAGCTAGAACGAGCGGCCCTAGGTTT
CGGAAGGGAGGATCAGGGATGTTGCGAGCGGCTGGAACCAGACGGTCCGATAGAGGAAGC
GGGCTCCATGGCTGCCCTCTGCTGCTGCCCTGCTGCTGTTGCTACCGCTGCTGCTGA
AGCTACACCTCTGCCGCAGTTGCGCTGGCTTCCGGGAGCTGGCCTTGCCTGCGAGCT
CTGTGCTGCAAAGGGCTCTCGAGCTCGCCCTGCCGCCGCTGCCGCCGACCCGGAAGG
TCCCGAGGGGGCTGCAGCTGCCCTGGCGCTCGGAACTGGCCAGCAGCGGCCGCG
ACACCTTCTCATTCACGGCTCGGGCGCTTAGCTACTCAGAGGCGAGCGAGAGTAAC
AGGGCTGCACGCCCTCCTACGTGCGTAGGCTGGACTGGGACCCGACGGCGCGACAG
CGCGAGGGGAGCGCTGGAGAAGCGAGCGGCCAGGCCGGAGCCGAGATGCAGCGCCG
GAAGCGGCCGGAGTTGCCGGAGGGGACGGTGCCTGCCAGAGGTGGAGGAGCCGCCG
CTGTCACCTGGAGCAACTGTGGCGTGTCTCCCCCTGGCCAGAGTTCTGTGGCTCTG
GTTGGGCTGCCAAGGCCCTGCGACTGCCCTGTGCCAGCTGGCCACCGCCCTGCCGGGCC
CCCTGCTGCACTGCCCTCGCAGCTGCCCGCGCGCGCTGGTGCCTGCCAGAGTTCTG
GAGTCCCTGGAGCCGGACCTGCCGCCCTGAGAGCCATGGGCTCCACCTGTGGCTGCAGG
CCAGGAACCCACCCCTGTTGAATTAGCGATTGCTGGCTGAAGTGTCCGCTGAAGTGGATG
GCCAGTGCCAGGATAACCTCTCTCCCCCAGAGCATAACAGACACGTGCCTGTACATCTC
ACCTCTGGCACCACGGCCTCCCCAAGGCTGCTGGATCAGTCATCTGAAGATCCTGCAATG
CCAGGGCTCTATCAGCTGTGTGGTCCACCAGGAAGATGTGATCTACCTGCCCTCCCAC
TCTACCACATGTCCGGTCCCTGCTGGCATCGTGGCTGCATGGCATTGGGCCACAGTG
GTGCTGAAATCCAAGTTCTCGGCTGGTCAGTTCTGGGAAGATTGCCAGCACAGGGTGAC
GGTGTCCAGTACATTGGGAGCTGTGCCGATACCTGTCAACCAGCCCCGAGCAAGGCAG
AACGTGCCATAAGGTCCGGCTGGCAGTGGCAGCAGGGCTGCCAGATACTGGAGCGT
TTTGTGCCGCTCGGGCCCTGCAGGTGCTGGAGACATATGGACTGACAGAGGGCACAGT
GCCACCATCAACTACACAGGACAGCGGGCGCTGTGGGGCGTCTGGCTTACAAGC
ATATCTTCCCTTCTCCTGATTGCTATGATGTCAACCAGGAGAGCCAATTGGGACCCC
CAGGGCACTGTATGCCACATCTCAGGTGAGCCAGGGCTGCTGGTGGCCCGTAAGCCA
GCAGTCCCCATTCTGGCTATGCTGGCGGCCAGAGCTGGCCAGGGGAAGTTGCTAAAGG
ATGTCTTCCGGCTGGGATGTTTCTCAACACTGGGACCTGCTGGTCTGCGATGACCAA
GGTTTCTCCGCTTCCATGATCGTACTGGAGACACCTTCAGGTGGAAGGGGAGAATGTGGC
CACAACCAGGTGGCAGAGGTCTCGAGGCCCTAGATTTCTCAGGAGGTGAACGTCTATG
GAGTCAGTGTGCCAGGGCATGAAGGCAGGGCTGGAATGGCAGCCCTAGTTCTGCGTCCCCC
CACGCTTGGACCTTATGCACTACACCCACGTGCTGAGAACCTGCCACCTTATGCCG
GCCCGGATCCTCAGGCTCCAGGAGTCTTGGCCACCACAGAGACCTCAAAACAGCAGAAAG
TTCGGATGGCAAATGAGGGCTTCGACCCAGCACCCACTGTACGTTCTGGAC
CAGGCTGTAGGTGCCTACCTGCCCTACAACATGCCGGTACAGCGCCCTCTGGCAGGAAA
CCTTCGAATTGAAACTTCCACACCTGAGGCACCTGAGAGAGGAACCTGTGGGTGGGG
CCGTTGCAGGTGTACTGGCTGTCAGGGATCTTTCTATACCAGAACCTGCCACTATTT
GTAATAAAATGTGGCTGGAGCTGATCCAGCTGTCTGACCTAAAAAAA
AAAAAAAAAGGGCGGCCGCACTCTAGAGTCGACCTGCAGTAGGGATAACAGGGTAATAAGC
TTGGCCGCCATGGCCCACTTGTATTGCAG

FIGURE 39

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50913
><subunit 1 of 1, 730 aa, 1 stop
><MW: 78644, pI: 7.65, NX(S/T): 2
MGVCQRTRAPWKEKSQQLERAALGFRKGSGMFASGWNQTVPIEEAGSMAALLLPLLLLLPL
LLLKLHLWPQLRWLPADLAFAVRALCCKRALRALARALAAAADPEGPEGGCSLAWRLAELAQQ
RAAHTFLIHSRRFSYSEAERESNRAARAFLRALGWDWGPDDGSGEGSAGEGERAAPGAGD
AAAGSGAEFAGGDGAARGGAAAPLSPGATVALLPAGPEFLWLWFGGLAKAGLRTAFVPTAL
RRGPLLHCLRSCGARALVLAPEFLESLEPDLPALRAMGLHLWAAGPGTHPAGISDLLAEVSA
EVDPVPVPGYLSSPQSITDTCLYIFTSGTGLPKAARISHLKILQCQGFYQLCGVHQEDVIYL
ALPLYHMSGSLLGIVGCMGIGATVVLKSKFSAGQFWEDCQQHRVTVFQYIGELCRYLVNQPP
SKAERGHKVRLAVGSGLRPDTWERFVRRFGPLQVLETYGLTEGNVATINYTGQRGAVGRASW
LYKHIFPFSLIRYDVTGEPIRDPQGHCMATSPGEPLLVAAPVSQQSPFLGYAGGPELAQGK
LLKDVFRPGDVFFNTGDLLVCDDQGFLRFHDRTGDTFRWKGENVATTEVAEVFEALDFLQEV
NVYGVTVPGHEGRAGMAALVLRPPHALDLMQLYTHVSENLPYARPRFLRLQESLATTEFK
QQKVRMANEGFDPSLSDPLYVLDQAVGAYLPLTTARYSALLAGNLRI
```

Type II transmembrane domain:

amino acids 45-65

Other transmembrane domain:

amino acids 379-398

cAMP- and cGMP-dependent protein kinase phosphorylation site
starting at amino acid 136

CUB domain protein motif

amino acids 254-261

putative AMP-binding domain siganture

amino acids 332-343

N-glycosylation sites

amino acids 37-40 and 483-486

FIGURE 40

CCTGTGTTAAGCTGAGGTTCCCTAGATCTGTATATCCCCAACACACATACCTCCACGCACA
CACATCCCCAAGAACCTCGAGCTCACACCAACAGACACACCGCGCATACACACTCGCTCTC
GCTTGTCCATCTCCCTCCCCGGGGAGCCGGCGCGCTCCACCTTGCAGCTCGGACACTCCGGC
GAGCCGAGCCCGCAGCGCTCCAGGATTCTCGGGCTCGAACCTCGGATTGCAGCTCTGAACCC
CCATGGTGGTTTTAACACTTCTTTCTCTCGTTGATTGCACCGTTCCA
TCTGGGGCTAGAGGAGCAAGGCAGCAGCCTCCCAGCCAGCCCTTGGCTGCCATCGT
CCATCTGGCTTATAAAAGTTGCTGAGCGCAGTCCAGAGGGCTGCGCTGCTCGTCCCCTCGG
CTGGCAGAAGGGGTGACGCTGGCAGCGCGAGGAGCGCGCCGCTGCCCTGGCGGGCTT
CGGCTTGAGGGCAAGGTGAAGAGCGCACCGCCGTGGGTTACCGAGCTGGATTGTATG
TTGCACCA**T**GCTTCTGGATCGGGCTGTGATTCTTCCCCTTGAGGCTGCTCTCCC
TCCCCGCCGGGCGGATGTGAAGGCTCGGAGCTGCGGAGAGGTCCGCCAGGCGTACGGTGC
AAGGGATTCAAGCCTGGCGGACATCCCCTACCAAGGAGATCGCAGGGAAACACTTAAGAATCTG
TCCTCAGGAATATACTGCTGCACCACAGAAATGGAAGACAAAGTTAAGCCAACAAAGCAAAC
TCGAATTGAAAACCTGTGGAAGAGACAAGCATTGTGCGCACCACTTTGTGTCAGG
CATAGAAATTGACGAATTTCGAGAGCTCCTGGAGAATGCAAGAAAAGTCACTAAATGA
TATGTTGTACGGACCTATGGCATGCTGTACATGCAAGAATTGAGCTTCCAGGACCTCT
TCACAGAGCTGAAAAGGTACTACACTGGGTAATGTGAATCTGGAGGAAATGCTCAATGAC
TTTGGGCTCGGCTCCTGGAACGGATGTTTCAGCTGATAAACCCCTCAGTATCACTCAGTGA
AGACTACCTGGAATGTGTGAGCAAATACACTGACCAGCTCAAGCCATTGGAGACGTGCC
GGAAACTGAAGATTCAAGGTTACCGCGCCTTCATTGCTGCCAGGACCTTGTCCAGGGCTG
ACTGTGGCAGAGAAAGTTGCAAACCGAGTTCCAAGGTCAAGCCAACCCCAGGGTGTATCCG
TGCCTCATGAAGATGCTGACTGCCCATCTGTCGGGCTTCCCAGTGTGAGGCCCTGCA
ACAACACTGTCTAACGTATGAAGGGCTGTTGGCAAATCAGGCTGACCTCGACACAGAG
TGAATCTGTTATAGATGCAATGCTCTTGGTGGCAGAGCAGTGGAGGGCATTCAACAT
TGAGTCGGTCATGGACCCGATAGATGTCAGGATTCTGAAGCCATTATGAACATGCAAGAAA
ACAGCATGCAGGTGCTGCAAAGGTCTTCAGGGATGTGGTCAGCCAAACCTGCTCCAGCC
CTCAGATCTGCCGCTCAGCTCTGAAAATTAAATACACGTTCAAGGCCCTACAATCTGA
GGAAAGACCAACAACTGCTGCAGGCACAAGCTGGACCGCTGGTCACAGACATAAAAGAGA
AATTGAAGCTCTCTAAAAGGTCTGGTCAGCATTACCTACACTATGCAAGGACGAGAGC
GTGACAGCGGGCACGTCAAACGAGGAGGAATGCTGGAACGGCACAGCAAAGCCAGATACTT
GCCTGAGATCATGAATGATGGCTCACCAACCAGATCAACAAATCCGAGGTGGATGTGACA
TCACTCGGCTGACACTTTCATCAGACAGCAGATTATGGCTCTCGTGTGATGACCAACAAA
CTAAAAAAAGCCTACAATGCAATGATGTCAATTCCAGGACACAAGTGATGAATCCAGTGG
CTCAGGGAGTGGCAGTGGTGCATGGATGACGTGTGCTTCAAGGAGTTGAGTTGTCA
CAGAGGCCCGCAGTGGATCCGACCGGAGAGAGGTGGACTCTTCTGCAGCCAGCGTGGC
CACTCCCTGCTCTGGTCTCTCACCTGCATTGCTGGACTGCAGAGACTGTGAG**A**
ATCTTGGTTTTGGTCAGATGAAACTGCATTAGCTATGCAATGCCAACTCACTTCTT
TTCTTACACTCTGGACAATGGACCATGCCACAAAACCTACCGTTCTATGAGAAGAGAG
CAGTAATGCAATCTGCCTCCCTTTGTTCCAAAGAGTACCGGGTGCCAGACTGAAC
CTTCCTCTTCAGCTATCTGTGGGACCTGTTATTCTAGAGAGAATTCTTACTCAA
ATTTTCGTACCAGGAGATTTCTTACCTTCATTGCTTTATGCTGCAGAAGTAAAGGAAT
CTCACGTTGTGAGGGTTTTCTCATTAAAAT

FIGURE 41

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50914
><subunit 1 of 1, 555 aa, 1 stop
><MW: 62736, pI: 5.36, NX(S/T): 0
MPSWIGAVILPLLGLLSSLPAGADVAKARSCGEVRQAYGAKGFSLADI PYQEIA GEHLRICPQ
EYTCCTTEMEDKLSQQSKLEFENLVEETSHFVRTTFVSRHKKFDEFFRELLENAEKS LNDMF
VRTYGM LYMQNSEVFQDLFTELKRYYTGGVNLEEMLNDFWARLLERMFQLINPQYHFSEDY
LECVSKYTDQLKPFGDVPRKLKIQVTRA FIAARTFVQGLTVGREVANRVSKVSP TPGCIRAL
MKMLYCPYCRGLPTVRPCNNYCLNV MKGCLANQADLDTEWNLFIDAMLLVAERLEGPFNIES
VMDPIDVKISEAIMNMQENS MQVSAKVFQGCGQPKPAPALRSARSAPENFNTRFRPYNPEER
PTTAAGTSLDRLVTDIKEKLKLSKKVWSALPYTICKDESVTAGTSNEECWNGHSKARYLPE
IMNDGLTNQINNPEVDVDITRPDTFIRQQIMALRVMTNKLKNAYNGNDVNFQDTSDESSGSG
SGSGCMDDVCPTEFEFVTTEAPAVDPDRREVDSSAAQRGHSLLSW LTCIVLALQRLCR
```

Signal peptide:

amino acids 1-23

FIGURE 4A

CGGACGCGTGGCGGACGCGTGGCAAAAGAACTCGGAGTGCCAAGCTAAATAAGTTAGCTGAGAAAACGCACG
CAGTTGCAGCGCTCGCCGGGTGCGCCAACACGCAAAGACCAAGCGGGCTCCGCGCGACCAGGCCGCGGGC
TAGGGACCCGGCTTGGCCTTCAGGCTCCCTAGCAGCGGGAAAAGGAATTGCTGCCGGAGTTCTCGGGAGGT
GGAGGGAGATCAGGAAACGGCTTCTCCTCACTTCGCCCTGGTAGTGTGGAGATTGGCAAACGCCCTAGG
AAAGGACTGGGAAAATAGCCCTGGAAAGTGGAGAAGGTGATCAGGAGGCCGTCAACTACGGCAGTTATCTG
TCTGATCAGAGCCAGACGCGACCGTCCACTCGCAGTTCTCCAGGTGAGTGTGGGACCGCAGGACAGACGCCGA
TCCCGCCGCCCTCCGTACCAGCACTCCAGGAGAGTCAGCCTCGCTCCCCAACGTCGAGGGCCTGGCACGA
AAAGTTCTGTCCACTGTGATTCTCAATTCTTGCTTGGTTTTCTCCAGAGAACTTTGGGTGGAGATATTA
ACTTTTTCTTTTTCTTGCTGGAGCTGCTCTAGGGAGGGGGAGGAGGAGGAGAAGTGAATGTGC
TGGAGAAGAGCGAGCCCTCCTGTTCTCCGGAGTCCCATTAAGCCATCCTCTGGAGATTAAAGTTGT
CGGACATGGTACAGCTGAGAGGAGAGGAGGAGTTCTGCCAGGTGGAGAGTCCTCACCTGGACCTGAGTGAATGGCTCCA
GGGGCTGTGGGGGATCCGCCTCCGCCTCTCCACAGGCCTGTGTCTGCCCTGGAAAGATGCTAGCAATGGGG
CGCTGGCAGGATTCTGGATCCTCTGCCCTCCTACTTATGGTACCTGTCTGCCCTGGCAGGCTTAGAAGAGGAGG
AAGAAGGGCCTACTAGCTCAAGCTGGAGAGAAACTAGAGCCCAGCACAACCTCCACCTCCAGCCCCATCTCA
TTTCATCCTAGCGGATGATCAGGGATTTAGAGATGTGGTTACACGGATCTGAGATTAAAACACCTACTCTG
ACAAGCTCGCTGCCAAGGAGTTAAACTGGAGAACTACTATGTCCAGCCTATTGCACACCATCCAGGAGTCAGT
TTATTACTGGAAAGTATCAGATACACACCGGACTTCAACATTCTATCATAAGACCTACCCACCCACTGTTAC
CTCTGGACAATGCCACCCCTACCTCAGAAACTGAAGGGAGTTGGATATTCAACGCATATGGTCGAAAATGGCACT
TGGGTTTAACAGAAAAGAATGCATGCCACCAGAAGAGGATTGATACCTTTGGTCCCTTTGGAAAGTG
GGGATTACTATACACACTACAAATGTGACAGTCCTGGATGTGGCTATGACTGTATGAAAACGACAATGCTG
CCTGGACTATGACAATGGCATATACTCCACACAGATGTACACTCAGAGAGTACAGCAAATCTAGCTCCATA
ACCCACAAAGCCTATATTTATATACTGCCTATCAAGCTGTTCAACACTGCTGCCATGCTTCTGCTTAGATGAAGCA
TCGAACACTACCGATCCATTATCAACATAAACAGGAGAAGATATGCTGCCATGCTTCTGCTTAGATGAAGCA
TCAACAACGTGACATTGGCTCTAAAGACTTATGGTTCTATAACAAACAGCATTATCATTACTCTTCAGATAATG
GTGGCCAGCCTACGGCAGGAGGAGTAACCTGGCTCTCAGAGGTAGCAAAGGAACATATTGGAAAGGAGGATCC
GGGCTGTAGGCTTGTGCATAGCCCACCTCTGAAAAACAAAGGAACAGTGTGTAAGGAACCTGTGCACATCACTG
ACTGGTACCCACTCTCATTCACTGGCTGAAGGACAGATTGATGAGGACATTCAACTAGATGGCTATGATATCT
GGGAGACCATAAGTGAGGGTCTCGCTCACCCGAGTAGATATTTCATACATTGACCCCTATACACCAAGGC
AAAAAATGGCTCTGGCAGCAGGCTATGGATCTGAAACACTGCAATCCAGTCAGCCATCAGAGTGCAGCACTG
GAAATTGCTTACAGGAAATCCTGGCTACAGCGACTGGTCCCCCTCAGTCTTCAGCAACCTGGACCGAACCG
GTGGCACAATGAACGGATCACCTGTCAACTGGAAAAGTGTATGGCTTCAACATCACAGCCACCCATATGA
GAGGGTGGACCTATCTAACAGGTATCCAGGAATCGTGAAGAAGCTCCTACGGAGGCTCTCACAGTTCAACAAAAC
TGCAGTGCCGGTCAGGTATCCCCCAAAGACCCAGAAGTAACCCCTAGGCTCAATGGAGGGTCTGGGACCATG
GTATAAAGAGGAAACCAAGAAAAGAAGCCAAGAAAAATCAGGCTGAGAAAAGCAAAGAAAAGCAAAAAAA
GAAGAAGAAACAGCAGAAAGCAGTCTCAGGTAACCCAGCAAATTGGCTCGATAATATGCTGGCTAAGCGTCA
GGCTTGTTCATGCTGTGCCACTCCAGAGACTCTGCCACCTGGCGCCACACTGAAAACGTCTGCTCAGTG
CCAAGGTGCTACTCTGCAAGCCACACTTAGAGAGAGTGGAGATGTTATTCTCGCTCCTTAGAAAACGTG
GTGAGTCCTGAGTTCACTGCTGTGCTTCAGTCAGTACACTGACCAAACACTGCTTGAATTATAGGAGGAGAACATA
ACCTACCATCCGCAAGCATGCTAATTGATGGAAGTTACAGGGTAGCATGATTAAACTACCTTGATAAAATTAC

FIGURE 42B

AGTCAAAGATTGTGTCACCTCAAAGGCCTTGAAGAATATATTTCTTGGTGAATTTGTATGTCTGTATATGA
CACTTGGGTTTTAATTAAATTCTATTTATATATAAATATGTTCTTCCTGTGAAAGCTGTTTCT
CACATGTGAACAGCTGCACCTCATTACCATGCGTGAGGGAAATGGCAAATAAGAATGTTGAGCACACTGCC
ACAATGAATGTAACATTTCTAAACACTTACTAGAAGAACATTCACTGAAATGACTTTATTTCTGCATACCATTAGAAGA
GAAAAATATTTGTTGTTTATAAAAAGTTATGCAAATGACTTTATTTCTGCATACCATTAGAAGA
ATTTATTCATTCTCAAATTATCAAGCACTGTAACACTATAAATTAAATGTAACACTGTGTGAATTCA
TAAAAAACATCATTCACTGAAACTTATAATCGTCAATTGTCATCAAGATTTGAATGTAATAAGATGAATATAT
ATTACTGGAAATTCAATGTTGTCAGAGTTGAGACAACCTTATTGTTCTATCATAAAACTATTTATGTATCTT
AATTATTAATGATTACTTATGGCACTAGAAAATTACTGTTGCTTCTGATCTAACTCTAGCTAAATT
GTATCATTGGCCTAAAAAATAAAAACTTACTAATAGGCAATTGAAGGAATGGTTGCTAACACCACAGTAA
TATAATATGATTTACAGATAGATGCTTCCCTGGCTATGACATGGAGAAAGATTTCCATAATAAACTAA
TATTTATATTAGGTTGGTGCAAAACTAGTTGCGGTTTCCATTAAAGTAATAACCTTACTCTTACAAAGT
GGACACTGTGGGAGATACAGAGAAATGGAAGATAACGGATCCTGCCTGGAGTAGGTAACCTGCTTGGAAACCC
ACATGCAAACGTATGAGGAGAATTAAAGGAGTATTATCAGTAATGAAGTTATCATGGGTATCAATGAGCATA
GATTGGTGTGGATCCTGTAGACCCCTGGTGTTCCTTGAAGTGCCTCTCCTAAATGCAGAGGCCTTGAAGCTTAC
AGTATACACTGAAAAGTCACAGATAGCTAGAATTATGATCTTGAAGTTAACTGTGATCTGAAATGTGT
GGTGGTATGACAGCATACCATTAAATACATTACATCACAGCTCAAAGGACTGTGATATAATCCATTATAC
AACTCAAAGGACTGTGATATAATCCATTATACAGCTCACAGTTCTGAAAATGTATAAAAGAATCTATAAT
CTAGTACTGAAATTACTAAATTGGTAAGATGATTTAAATGATTTAATTAAACATTATTTCTAGAATATAT
GGCTCCATTATTTATTTATAGTGTAAAGTTGTATTTCTAAAGTTGTGTTGTCGACAGTATCTTAAATGAG
TCTTAAAAATAAGGCATATTGTTCATGTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AA

FIGURE 43

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss:DNA48296
><subunit 1 of 1, 515 aa, 1 stop
><MW: 56885, pI: 6.49, NX(S/T): 5
MAPRGCAHGPPPPSPQACVCPGKMLAMGALAGFWILCLLTYGYLSWGQALEEEEGLALLAQA
GEKLEPSTTSTSQPHLIFILADDQGFRDVGYHGSEIKPTLDKLAEGVKLENYVQPICTP
SRSQFITGKYQIHTGLQHSIIRPTQPNCLPLDNATLPQKLKEVGYSTHMVGKWHLGFRKEC
MPTRRGFDTFFGSSLGSGDYYTHYKCDSPGMCGYDLYENDNAAWDYDNGIYSTQMYTQRVQQ
ILASHNPTKPIFLYTAYQAVHSPLQAPGRYFEHYRSIININRRRYAAMLSCLDEAINVTLA
LKTYGFYNNSSIIYSSDNGGQPTAGGSNWPLRGSKGTWEGGIRAVGFVHSPLLKNKGTVCK
ELVHITDWYPTLISLAEGQIDEDIQLDGYDIWETISEGLRSPRVDILHNIDPYTPRQKMAPG
QQAMGSGTLQSSQPSECSTGNCLQEILATATGSPLSLSATWDRTGGTMNGSPCQLAKVYGF
TSQPTHMRGWTYLTGIQES

Important Features:

Signal Peptide:

amino acids 1-37

Sulfatases signature 1.

amino acids 120-132

Sulfatases signature 2.

amino acids 168-177

Tyrosine kinase phosphorylation site.

amino acids 163-169

N-glycosylation sites.

amino acids 157-160, 306-309 and 318-321

FIGURE 44

CGGACGCGTGGGTGCGAGTGGAGCGGAGGACCCGAGCGGCTGAGGAGAGAGGGAGGAGGCGGCGGC
TTAGCTGCTACGGGGTCCGGCCGGCGCCCTCCGAGGGGGCTCAGGAGGAGGAAGGAGGAC
CCGTGCGAGAATGCCTCTGCCCTGGAGCCTGCGCTCCGCTGCTCTCCTGGTGGCAG
GTGGTTTGGGAACGCGGCCAGTCAAGGCATCACGGTTGTTAGCATCGGCACGTCAGCCT
GGGGTCTGTCACTATGGAACTAAACTGGCCTGCTGCTACGGCTGGAGAAGAAACAGCAAGGG
AGTCTGTGAAGCTACATGCGAACCTGGATGTAAGTTGGTAGTGCCTGGGACCAAACAAAT
GCAGATGCTTCCAGGATACACCGGGAAAACCTGCAGTCAGTCAGATGTGAATGAGTGTGGAATG
AAACCCCGGCCATGCCAACACAGATGTGTGAATACACACCGGAAGCTACAAGTGCTTTGCCT
CAGTGGCCACATGCTCATGCCAGATGCTACGTGTGAACTCTAGGACATGTGCCATGATAA
ACTGTCAGTACAGCTGTGAAGACACAGAAGAAGGGCACAGTGCCTGTGCTCCATCCTCAGGA
CTCCGCCTGGCCCAAATGGAAGAGACTGCTAGATATTGATGAATGTGCCTCTGGTAAAGT
CATCTGTCCTACAATCGAAGATGTGTGAACACATTGGAAGCTACTACTGCAAATGTCACA
TTGGTTTCAACTGCAATATATCAGTGGACGGATATGACTGTATAGATATAAATGAATGTA
ATGGATAGCCATACGTGCAGCCACCATGCCAATTGCTCAATACCCAAAGGGCTTCAAGTG
TAAATGCAAGCAGGGATATAAAGGCAATGGACTTCGGTGTCTGCTATCCCTGAAAATTCTG
TGAAGGAAGTCCTCAGAGCACCTGGTACCATCAAAGACAGAATCAAGAAGTTGCTTGCTCAC
AAAAACAGCATGAAAAGAAGGCAAAATTAAAATGTTACCCAGAACCCACCCAGGACTCC
TACCCCTAAGGTGAACCTGCGAGCCCTCAACTATGAAGAGATAGTTCCAGAGGGCGGGAACT
CTCATGGAGGTAAAAAGGGAATGAAGAGAAATGAAAGAGGGCTTGAGGATGAGAAAAGAG
AAGAGAAAGCCCTGAAGAATGACATAGAGGGAGCGAACGCTCGAGGAGATGTGTTTCCCT
AAGGTGAATGAAGCAGGTGAATTGGCCTGATTCTGGTCAAAGGAAAGCGCTAACCTCCAA
ACTGGAACATAAAGATTAAATATCTCGGTTGACTGCAGCTCAATCATGGGATCTGTGACT
GGAAACAGGATAGAGAAGATGATTTGACTGGAATCCTGCTGATCGAGATAATGCTATTGGC
TTCTATATGGCAGTTCCGGCTTGGCAGGTACAAGAAAGACATTGGCGATTGAAACTTCT
CCTACCTGACCTGCAACCCCAAAGCAACTCTGTTGCTTTGATTACCGGCTGGCCGGAG
ACAAAGTCGGAAACTTCGAGTGTGAAAACAGTAACAATGCCCTGGCATGGGAGAAG
ACCACGAGTGAGGATGAAAAGTGGAGACAGGGAAATTCAAGTGTATCAAGGAACGTG
TACCAAAAGCATCATTGGCAGGAACTAGCTGAAAATTGTAATGTACCAACAGAAATATT
TAAGATGCCTTCTTGTATAAGATATGCCAATATTGCTTAAATATCATACTGTATCT
TCTCAGTCATTCTGAATCTTCCNCATTATATTAAATNTGGAAANGTCAGTTATCTC
CCCTCCTCNGTATATCTGATTGTATANGTANGTTGATGNGCTCTCTACAACATTCTA
GAAAATAGAAAAAAAGCACAGAGAAATGTTAACTGTTGACTCTTATGATAACTTCTTGG
AACTATGACATCAAAGATAGACTTTGCCTAAGTGGCTTAGCTGGTCTTCATAGCCAAAC
TTGTATATTAAATTCTTGTATAATAA

FIGURE 45

MPLPWSLALPLLLSWVAGGFNGNAASARHHGLLASARQPGVCHYGTKLACCYGWRNSKGVC
ATCEPGCKFGECVGPNCRCFPGYTGKTCSDVNECGMKPRPCQHRCVNTHGSYKCFCLSGH
MLMPDATICNSRTCAMINQYSQEDTEEGPQCLCPSSGLRLAPNGRDLCLDIDEASGKVICP
YNRRCVNTFGSYYCKCHIGFELQYISGRYDCIDINECTMDSHTCSHHANCNTQGSFKCKCK
QGYKGNGLRCSAIPENSVKEVLRAPGTIKDRIKLLAHKNSMKKAKIKNVTPPEPTRTPPK
VNLQPFNYEEIVSRGGNSHGGKKGNEEK

Signal peptide:

amino acids 1-21

EGF-like domain cysteine pattern signature.

amino acids 80-91

Calcium-binding EGF-like domains

amino acids 103-124, 230-251 and 185-206

FIGURE 46

GGGAGCTGCTGCTGGCTGCTGGTCTGCGCGCTGCTCCTGCTCTGGTGCAGCTGCTG
CGCTTCCTGAGGGCTGACGGCGACCTGACGCTACTATGGCCGAGTGGCAGGGACGACGCC
AGAATGGGAGCTGACTGATATGGTGGTGGGTGACTGGAGCCTCGAGTGGAATTGGTGAGG
AGCTGGCTTACCAAGTTGTCTAAACTAGGAGTTCTCTGTGCTGTCAGCCAGAAGAGTCAT
GAGCTGGAAAGGGTGAAGAAGATGCCTAGAGAATGCAATTAAAAGAAAAAGATATACT
TGTGTTGCCCTTGACCTGACCGACACTGGTCCATGAAGCGGCTACCAAAGCTGTTCTCC
AGGAGTTGGTAGAATCGACATTCTGGTCAACAATGGTGAATGTCCAGCGTTCTGTGC
ATGGATACCAGCTGGATGTCTACAGAAAGCTAATAGAGCTTAACCTAGGGACGGTGTC
CTTGACAAAATGTGTTCTGCCCTCACATGATCGAGAGGAAGCAAGGAAAGATTGTTACTGTGA
ATAGCATCCTGGGTATCATATCTGTACCTCTTCCATTGGATACTGTGCTAGCAAGCATGCT
CTCCGGGGTTTTTAATGGCCTTCGAACAGAACTGCCACATACCCAGGTATAATAGTTTC
TAACATTGCCAGGACCTGTCAATCAAATATTGTGGAGAATTCCCTAGCTGGAGAAGTCA
CAAAGACTATAGGCAATAATGGAGACCAAGTCCCACAAGATGACAACCAGTCGTTGTGCGG
CTGATGTTAACATCAGCATGCCAATGATTGAAAGAAGTTGGATCTCAGAACACCTTCTT
GTTAGTAACATATTGTGGCAATACATGCCAACCTGGCCTGGGATAACCAACAAGATGG
GGAAGAAAAGGATTGAGAACCTTAAGAGTGGTGTGGATGCAGACTCTTCTTATTTAAAATC
TTAAGACAAAACATGACTGAAAAGAGCACCTGTACTTTCAAGCCACTGGAGGGAGAAATG
GAAAACATGAAAACAGCAATCTCTTATGCTTCTGAATAATCAAAGACTAATTGTGATT
ACTTTTAATAGATATGACTTGTGCTTCAACATGGAATGAAATAAAAATAATAAAAAG
ATTGCCATGAATCTGCAAA

FIGURE 47

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA36343
><subunit 1 of 1, 289 aa, 1 stop
><MW: 32268, pI: 9.21, NX(S/T): 0
MVVWVTGASSGIGEELAYQLSKLGVSLVLSARRVHELERVKRRCLENGNLKEKDILVLPPLDL
TDTGSHEAATKAVLQEFGRIDILVNNGGMSQRSLCMDTSLDVYRKLIENYLGTVSLTKCVL
PHMIERKQGKIVTVNSILGIISVPLSIGYCASKHALRGFFNGLRTELATYPGIIVSNICPGP
VQSNIVENSLAGEVTKTIGNNGDQSHKMTTSRCVRLMLISMANDLKEVWWISEQPFLVTYLW
QYMPPTWAWWITNKGKKRIENFKSGVDADSSYFKIFKTKHD
```

Important Features:

Signal Peptide:

amino acids 1-31

Transmembrane domain:

amino acids 136-157

Tyrosine kinase phosphorylation site.

106-113 and 107-114

Homologous region to Short-chain alcohol dehydrogenase

amino acids 80-90, 131-168, 1-13 and 176-185

FIGURE 48

GCGACGTGGCACC GCCATCAGCTGTT CGCGCGTCTTCTCCTCCAGGTGGGGCAGGGGTTTC
GGGCTGGTGGAGCATGTGCTGGGACAGGACAGCATCCTCAATCAATCAACAGCATATT CGG
TTGCATCTTCTACACACTACAGCTATTGTTAGGTTGCCTGCGGACACGCTGGGCCTCTGTCC
TGATGCTGCTGAGCTCCCTGGTGTCTCGCTGGTTCTGTCTACCTGGCCTGGATCCTGTTC
TTCGTGCTCTATGATTCTGCATTGTTGTATCACCACCTATGCTATCAACGTGAGGCTGAT
GTGGCTCAGTTCCGGAAGGTCCAAGAACCCCAGGGCAAGGCTAAGAGGGACTGAGCCCTCA
ACCCAAGCCAGGCTGACCTCATCTGCTTGTCTTCAAGCCGCTCAGCGTGCCTGTG
GACAGCGTGGCCCCGGCCCCAACGCTCAGGAGGGCAACACAGTCCTGGCGAGTGGCC
TGGCAGGCCAGTGTGAGGAGGCAAGGAGCCACATCTGCAGCGGCTCCCTGGTGGCAGACAC
CTGGGTCCACTGCTGCCACTGCTTGAAGAAGGCAGCAGCAACAGAACAGACTGAATTCTGGT
CAGTGGTCCCTGGTTCTCTGCAGCGTGAGGAGCTCAGCCCTGGGGCCAAGAGGGTGGGGTG
GCTGCCCTGCAGTTGCCAGGGCTATAACCAACTACAGCCAGGGCTCAGACCTGGCCCTGCT
GCAGCTCGCCCACCCCACGACCCACACACCCCTCTGCCTGCCAGGCCGCCCCATCGCTTCC
CCTTGGAGCCTCTGCTGGCCACTGGCTGGATCAGGACACCAAGTGTATGCTCTGGGACC
CTACGCAATCTGCCCTGCGTCTCATCAGTCGCCACATGTAAGTGTATCTACAACCAAGCT
GCACCAGCGACACCTGTCCAACCCGGCCCTGGGATGCTATGTGGGGCCCCAGCCTG
GGTGCAGGGCCCTGTCAAGGAGATTCCGGGGCCCTGTGCTGTGCTCGAGCCTGACGGA
CACTGGGTTCAAGGCTGGCATCATCAGCTTGCATCAAGCTGTGCCAGGAGGACGCTCCTGT
GCTGCTGACCAACACAGCTGCTCACAGTTCTGGCTGCAGGCTCGAGTTCAAGGGGAGCTT
TCCGGCCAGAGGCCAGAGACCCGGAGATGAGTGTAGGAGACAGCTGTGTAGCCTGTGGA
TCCTTGAGGACAGCAGGTCCCCAGGCAGGAGCACCCCTCCCCATGGCCCTGGGAGGCCAGGCT
GATGCACCAGGGACAGCTGCCCTGTGGCGGAGCCCTGGTGTCAAGAGGAGGCGGTGCTAAGT
CTGCCCACTGCTTATTGGCGCCAGGCCAGAGGAATGGAGCGTAGGGCTGGGACCAAGA
CCGGAGGAGTGGGCCTGAAGCAGCTCATCCTGCATGGAGCCTACACCCACCCCTGAGGGGG
CTACGACATGCCCTCCTGCTGGCCAGCCTGTGACACTGGGAGGCCAGCCTGCCGGCCCC
TCTGCCTGCCCTATCCTGACCACCATGCCCTGATGGGAGCGTGGCTGGGTTCTGGGACGG
GCCGCCAGGAGCAGGCATCAGCTCCCTCAGACAGTGCCCGTACCCCTCTGGGCTAG
GCCCTGCAGCCGGCTGCATGCAGCTCTGGGGTGATGGCAGCCCTATTCTGCCGGGATGG
TGTGTACCAAGTGTGGGTGAGCTGCCAGCTGTGAGGGCTGTCTGGGACCAACTGGTG
CATGAGGTGAGGGCACATGGTCTCTGGCCGGCTGCACAGCTCGGAGATGCTTGCAAGG
CCCCGCCAGGCCGGCGGTCTCACCGCGCTCCCTGCTATGAGGACTGGTCAGCAGTTGG
ACTGGCAGGTCTACTTCGCGAGGAACCAGAGCCGAGGCTGAGCCTGGAAGCTGCCCTGGCC
AACATAAGCCAACCAACCAGCTG**TGAC**AGGGGACCTGGCATTCTCAGGACAAGAGAAATGC
AGGCAGGCCAAATGGCATTACTGCCCTGTCTCCCCACCCCTGTCATGTGTGATTCCAGGCAC
CAGGGCAGGCCAGAAGCCAGCAGCTGTGGGAAGGAACCTGCCCTGGGACAGGTGCCA
CTCCCCACCCCTGCAGGACAGGGGTGTCTGGACACTCCCACCCAACTCTGCTACCAAGC
AGGCCTCTCAGCTTCCCTCTTACTCTTCAGATAACATCACGCCAGCCACGTTGTT
TGAAAATTCTTTTTGGGGGGCAGCAGTTTCCTTTAAACTAAATAATTGTTAC
AAAATAAAA

FIGURE 49

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA40571
MLLSSLVSLAGSVYLAWLFFVLYDFCIVCITYAINVSLMWLSFRKVQEPOGKAKRHGNTV
PGEWPWQASVRRQGAHICSGSLVADTWVLTAAHCFEKAATELNSWSVVLGSLQREGLSPGA
EEVGVAALQLPRAYNHYSQGSDLALLQLAHPTTHTPLCLPQPAHRFPFGASCWATGWDQDTS
DAPGTLRNRLRLISRPTCNCIYNQLHQRLHSNPAPGMLCGGPQPGVQGPCQGDGGPVLC
LEPDGHWVQAGIISFASSCAQEDAPVLLNTAAHSSWLQARVQGAAFLAQSPETPEMSDEDS
CVACGSLRTAGPQAGAPSPWPWEARLMHQQLACGGALVSEEAVLTAAHCFIGRQAPEEWSV
GLGTRPEEWGLKQLILHGAYTHPEGGYDMALLLAQPVTLGASLRPLCLPYPDHHPDGERG
WVLGRARPAGISSLQTVPTLLGPRACSRHLAAPGGDGSPILPGMVCTSAGELPSCEGLS
GAPLVHEVRGTWFLAGLHSFGDACQGPARPAVFTALPAYEDWVSSLDWQVYFAEEPEPEAEP
GSCLANISQPTSC
```

Important features:

Signal peptide:

amino acids 1-15

Homologous region to Serine proteases, trypsin family

amino acids 79-95, 343-359 and 237-247

N-glycosylation sites.

amino acids 37-40 and 564-567

Kringle domains

amino acids 79-96, 343-360 and 235-247

FIGURE 50

CGGGCCGCCCGCCCCGCCCCATTGGGCCGGCCTCGCTGCCGCGGCGACTGAGCCAGGCTGG
GCCGCGTCCCTGAGTCCCAGAGTCGGCGCGCGGGCAGGGCAGCCTCCACCACGGGGAG
CCCAGCTGTCAGCCGCCTCACAGGAAGATGCTGCGTCGGCGGGCAGCCCTGGCATGGTGT
GCATGTGGGTGCAGCCCTGGGAGCACTGTGGTTCTGCCTCACAGGAGCCCTGGAGGTCCAGG
TCCCTGAAGACCCAGTGGTGGCACTGGTGGCACCGATGCCACCCCTGTGCTGCTCCTCTCC
CCTGAGCCTGGCTTCAGCCTGGCACAGCTAACCTCATCTGGCAGCTGACAGATAACAAACA
GCTGGTGCACAGCTTGCTGAGGGCCAGGACCAGGGCAGCGCCTATGCCAACCGCACGGCCC
TCTTCCCACCTGCTGGCACAGGGCACAGCATTGAGGCTGCAGCGCGTGCAGCGCTGCG
GACGAGGGCAGCTCACCTGCTCGTGCAGCATCCGGGATTCGGCAGCGCTGCCGTAGCCT
GCAGGTGGCCGCTCCCTACTCGAACCCCAGCATGACCCCTGGAGGCCAACAAAGGACCTGCC
CAGGGACACGGTGACCATCACGTGCTCCAGCTACCAGGGTACCCCTGAGGCTGAGGTGTC
TGGCAGGATGGCAGGGTGTGCCCTGACTGGCAACGTGACCACGTGAGATGCCAACGA
GCAGGGCTTGTGATGTGCACAGCGTCTGCGGGTGGTGTGGGTGCGAATGGCACCTACA
GCTGCCTGGTGCACAGCGTCTGCGGGTGGTGTGGGTGCGAATGGCACCTACA
CAGCCTATGACATTCCCCCAGAGGCCCTGCGGGTGGTGTGGGTGCGAATGGCACCTACA
TGCACGTGCTGGTGGCCCTGGCTTCGTGTGGAGAAAGATCAAACAGAGCTGTGAGGAGG
AGAATGCAGGAGCTGAGGACCAGGATGGGAGGGAGAAGGCTCCAAGACAGCCCTGCAGCCT
CTGAAACACTCTGACAGCAAAGAAGATGGACAAGAAATAGCCTGACCATGAGGACCAGG
GAGCTGCTACCCCTCCCTACAGCTCCTACCCCTGGCTGCAATGGGCTGCACGTGAGCCC
TGCCCCAACAGATGCACCTGCTGACAGGTGGCTCTCCAAAGGATGCGATACAC
AGACCACTGTGCAGCCTATTCTCAATGGACATGATTCCAAGTCATCCTGCTGCCTTT
TTCTTATAGACACAATGAACAGACCACCCACAACCTTAGTTCTCTAAGTCATCCTGCTGCT
GCCTTATTCACAGTACATACATTCTTAGGGACACAGTACACTGACCATCACCACCC
TTCTTCCAGTGCCTGGGACCATCTGGCTGCCTTTCTCCAAAAGATGCAATATTAGA
CTGACTGACCCCTGCCTATTCAACAGACACGATGCATAGTCACCCGGCCTGTT
TCCAATGGCCGTGATAACTAGTGATCATGTTCAGCCCTGCTCCACCTGCATAGAATCTT
TCTTCTCAGACAGGGACAGTGCGGCCTCAACATCTCCTGGAGTCTAGAAGCTGTTCT
CCCTCCTCCCTGCCCAAGTGAAAGACAGGGCAGGGCCAGGAATGCTTGAGGACACCG
AGGGGACTGCCACCCACCCACCATGGTGTATTCTGGGCTGGGAGTCTAGAAGCTGTTCT
TTGCCTCTGGCCAGCTCCTGGCCTCTGGTAGAGTGAGACTCAGACGTTCTGATGCCTCC
GATGTCATCTCCCTGCCCAAGGAATGGAAGATGTGAGGACTTCTAATTAAATGTGGGAC
TCGGAGGGATTTGTAAACTGGGGTATATTGGGAAAATAATGTCTTGAAAAAAA
AAAAAAAAAAAAAA

FIGURE 51

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA41386
><subunit 1 of 1, 316 aa, 1 stop, 1 unknown
><MW: -1, pI: 4.62, NX(S/T): 4
MLRRRGSPGMGVHVGAAALGALWFCLTGALEVQVPEDPVVALVGTDATLCCSFSPFGSLAQ
LNLIWQLTDTKQLVHSFAEGQDQGSAYANRTALFPDLLAQGNASLRLQRVRVADEGSFTCFV
SIRDFGSAAVSLQVAAPYSKPSMTLEPNKDLRPGDTVTITCSSYQGYPEAEVFWQDGQGVPL
TGNVTTSQMANEQGLFDVHSVLRVVLGANGTYSCLVRNPVLQQDAHXSVTITGQPMTFPPEA
LWVTVGGLSVCLIAALLVALAFVCWRKIKQSCEEENAGAEDQDGEGEGSKTALQPLKHSDSKED
DGQEIA
```

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 251-270

N-glycosylation site.

amino acids 91-94, 104-107, 189-192 and 215-218

Homologous region to Immunoglobulins and MHC

amino acids 217-234

FIGURE 52

TTCTGTGACCCCTTGAGAAAAGAGTTGGTGGTAAATGTGCCACGTCTCTAAGAAGGGGGAGTC
CTGAACCTTGTCTGAAGCCCTTGTCCGTAAGCCTGAACACTACGTTCTTAAATCTATGAAGTCG
AGGGACCTTCGCTGCTTTGTAGGGACTCTTCCTGCTTCAGCAACATGAGGCTTTCT
TGTGGAACCGGGCTTGTGACTCTGTCGTCATTGATTGGGCTTGATCCCTGAACCA
GAAAGTAAAATTGAAGTTCTCCAGAAGCCATTCATCGCCATCGCAAGACCAAAGGAGGGGA
TTGATGTTGGTCCACTATGAAGGCTACTTAGAAAAGGACGGCTCCTTATTCACCCACTC
ACAAACATAACAATGGTCAGCCCATTGGTTACCCGGCATCCTGGAGGCTCTCAAAGGT
TGGGACCAAGGGCTTGAAAGGAATGTGTGAGGAGAGAAGAGAAAGCTCATCATTCCCTGC
TCTGGGCTATGGAAAAGAAGGAAAAGGTAAAATTCCCCCAGAAAGTACACTGATATTAAATA
TTGATCTCCTGGAGATTGAAATGGACCAAGATCCCATGAATCATTCCAAGAAATGGATCTT
AATGATGACTGGAAACTCTCAAAGATGAGGTTAACGATATTAAAGAAGGAGTTGAAAAA
ACATGGTGCCTGGTGAATGAAAGTCATCATGATGCTTGGTGGAGGATATTTGATAAAAG
AAGATGAAGACAAAGATGGGTTATATCTGCCAGAGAATTACATATAAACACGATGAGTTA
TAGAGATACATCTACCCTTTAATATAGCACTCATCTTCAAGAGAGGGCAGTCATCTTAA
AGAACATTATTTATACAAATGTTCTTCTGCTTGTGTTTATTTATATATTTTT
CTGACTCCTATTTAAAGAACCCCTTAGGTTCTAAGTACCCATTCTTCTGATAAGTTATT
GGGAAGAAAAGCTAATTGGTCTTGAATAGAAGACTCTGGACAATTTCACCTTCACAG
ATATGAAGCTTGTGTTACTTCTCACTTATAAATTAAAATGTTGCAACTGGGAATATACC
ACGACATGAGACCAGGTTAGCACAATTAGCACCCATTCTGCTTCCCTATTTC
TCCAAGTTAGAGGTCAACATTGAAAAGCCTTGTCAATAGCCAAGGCTGCTATTTCAT
GTATAATGAAATAGTTATGTGTAACGGCTCTGAGTCTGCTTGAGGACCAGAGGAAAAA
TGGTTGTTGGACCTGACTTGTAAATGGCTACTGCTTACTAAGGAGATGTGCAATGCTGAAG
TTAGAAACAAGGTTAATAGCCAGGCATGGTGGCTATGCCTGTAATCCAGCAGTGGAG
GCTGAGGCGGGCGGATCACCTGAGGTTGGAGTTGAGACCAGCCTGACCAACACGGAGAAA
CCCTATCTCTACTAAAATACAAAGTAGCCCGGCGTGGTGTGCGTGCCTGTAATCCAGCT
ACCCAGGAAGGCTGAGGCAGAACACTTGAACCCAGGCGAGGTTGCGGTAAGCCGAG
ATCACCTNCAGCCTGGACACTCTGTCCTGAAAAAAAGAAAAGAACACGGTTAATACCATATNA
ATATGTATGCATTGAGACATGCTACCTAGGACTTAAGCTGATGAAGCTTGGCTCTAGTGAT
TGGTGGCCTATTATGATAAAATAGGACAAATCATTATGTGAGTTCTTGTAATAAAATG
TATCAATATGTTAGATGAGGTAGAAAGTTATATTATATTCAATATTACTTCTTAAGGC
TAGCGGAATATCCTCCTGGTTCTTAATGGGTAGTCTATAGTATATTACTACAATAACA
TTGTATCATAAGATAAAAGTAGTAAACCAAGCTACATTCTCCATTCTGTCATCAAAAAC
TGAAGTTAGCTGGGTGTGGTGGCTATGCCTGTAATCCAGCAGTCTGGGGCCAAGGAGGG
TGGATCAGTGGAGATCAGGAGTTCAAGACCAAGCAGCCTGGCCAACATGGTAAACCTTGTCTCA
CTAAAAATACAAAAATTAGCCAGGCGTGGTGGTCACACCTGTAGTCCCAGCTACTCAGGAG
GCTGAGACAGGAGATTGCTGAAACCCGGGAGGCGGAGGTTGCAAGTCAAGATTGTGCC
ACTGCACTCCAGCCTGGGTGACAGAGCAAGACTCCATCTCAAAAAAAAAAGAACAGCAGA
CCTACAGCAGCTACTATTGAATAAAACCTATCCTGGATTTT

FIGURE 53

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44194
><subunit 1 of 1, 211 aa, 1 stop
><MW: 24172, pI: 5.99, NX(S/T): 1
MRLFLWNAVLTLFVTSLIGALIPEPEVKIEVLQKPFICHRKTGGDMLVHYEGYLEKD GSL
FHSTHKHNNNGQPIWFTLGLILEALKGWDQGLKGMCVGEKRKLIIPPALGYGKEGKGKIPPEST
LIFNIDLLEIRNGPRSHESFQEMDLNDDWKLSKDEVKAYLKKEFEKGAVVNESH DALVED
IFDKEDEDKDGFISAREFTYKHDEL
```

Important features:

Signal peptide:

amino acids 1-20

N-glycosylation site.

amino acids 176-179

Casein kinase II phosphorylation site.

amino acids 143-146, 156-159, 178-181 and 200-203

Endoplasmic reticulum targeting sequence.

amino acids 208-211

FKBP-type peptidyl-prolyl cis-trans isomerase

amino acids 78-114 and 118-131

EF-hand calcium-binding domain.

amino acids 191-203, 184-203 and 140-159

S-100/ICaBP type calcium binding domain

amino acids 183-203

FIGURE 54

AATAAAGCTCCTTAATGTTGATATGTCCTTGAAGTACATCCGTGCATTTTTTTAGCAT
CCAACCATTCCCTCCCTGTAGTTCTGCCCTCAAATCACCTCTCCGTAGCCCACCCGA
CTAACATCTCAGTCTCTGAAAATGCACAGAGATGCCCTGGCTACCTGCCCTGCCCTCAGCCT
CACGGGGCTCAGTCTCTTTCTCTTGGTGCACAGGACGGAGCATGGAGGTACAGTAC
CTGCCACCCTCAACGTCTCAATGGCTCTGACGCCCTGCCCTGCACCTTCAACTCCTGC
TACACAGTGAACCAACAGTTCTCCCTGAACGGACTTACCAAGGAGTGCAACAACGTGCTC
TGAGGAGATGTTCCCTCCAGTTCCGCATGAAGATCATTAACCTGAAGCTGGAGCGGTTCAAG
ACCGCGTGGAGTTCTCAGGGAACCCCAGCAAGTACGATGTGCGGTGATGCTGAGAACCGTG
CAGCCGGAGGATGAGGGGATTACAACACTGCTACATCATGAACCCCCCTGAGCGGGACTCCACGGTGG
CCATGGCAAGATCCATCTGCAGGTCTCATGGAAAGAGCCCCCTGAGCGGGACTCCACGGTGG
CCGTGATTGTGGGTGCCTCCGTCGGGGCTCCTGGCTGTGGTCATCTTGGTGCTGATGGTG
GTCAAGTGTGTGAGGAGAAAAAAAGAGCAGAACGACTGAGCACAGATGACCTGAAGACCGAGGA
GGAGGGCAAGACGGACGGTGAAGGCAACCCGATGATGGGCCAAGTAGTGGGTGGCCGGCC
CTGCAGCCTCCCGTGTCCCGTCTCCTCCCCCTCCGCCCTGTACAGTGACCCCTGCCCTGCTCG
CTCTTGGTGTGCTTCCGTGACCTAGGACCCCAGGGCCCACCTGGGGCTCCTGAACCCCCG
ACTTCGTATCTCCACCCCTGCACCAAGAGTGACCCACTCTTCCATCCGAGAAACCTGCCA
TGCTCTGGGACGTGTGGGCCCTGGGAGAGAGAGAAAGGGCTCCACCTGCCAGTCCCTGG
GGGGAGGCAGGAGGCACATGTGAGGGTCCCAGAGAGAACGGAGTGGGTGGCAGGGGTAGA
GGAGGGCCGCTGTCACCTGCCAGTGCTGCCTGGCAGTGGCTTCAGAGAGGACCTGGTGG
GGAGGGAGGGCTTCCTGTGCTGACAGCGCTCCCTCAGGAGGGCCTGGCCTGGCACGGCTG
TGCTCCTCCCTGCTCCCAGCCCAGAGCAGCCATCAGGCTGGAGGTGACGATGAGTTCTGA
AACTTGGAGGGCATGTTAAAGGGATGACTGTGCATTCCAGGGCACTGACGGAAAGCCAGGG
CTGCAGGCAAAGCTGGACATGTGCCCTGGCCAGGAGGCCATGTTGGGCCCTGTTCCATT
GCTAGTGGCCTCCTGGGCTCCTGGCTCTTAATCCCTAGGACTGTGGATGAGGCCAG
ACTGGAAGAGCAGCTCCAGGTAGGGGCCATGTTCCCAGCGGGGACCCACCAACAGAGGCC
AGTTTCAAAGTCAGCTGAGGGCTGAGGGTGGGCTCCATGGTGAATGCAGGTTGCTGCAG
GCTCTGCCCTCTCCATGGGTAACCACCCCTGCCTGGCAGGGCAGCCAAGGCTGGAAAT
GAGGAGGCCATGCACAGGGTGGGCCAGTTCTTGGGCTTCAGTGAGAACCTCCAGTT
GCCCTTGGTGGGTTCCACCTGGCTTTGGCTACAGAGAGGGAAAGGCTGAGGCC
GCATAAGGGGAGGCCTTGGAACCTGAGCTGCCAATGCCAGCCCTGTCCCATCTGGGCCACG
CTACTCGCTCCTCTCCAAACAACCTCCCTCGTGGGACAAAAGTGACAATTGTAAGGCCAGGC
ACAGTGGCTCACGCCGTAAATCCCAGCACTTGGGAGGCCAAGGCGGGTGGATTACCTCCAT
CTGTTTAGAGAAATGGGAAAACCCATCTACTAAAATACAAGAATTAGCTGGCGTG
GTGGCGTGTGCTGTAATCCCAGCTATTGGGAGGCTGAGGCAGGAGAATCGCTTGAGCCCG
GGAAGCAGAGGTTGCAGTGAACGTGAGATAGTGTAGTGCCACTGCAATTAGCCTGGGTGAC
ATAGAGAGACTCCATCTAAAAAA

FIGURE 55

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45415
<subunit 1 of 1, 215 aa, 1 stop
<MW: 24326, pI: 6.32, NX(S/T): 4
MHRDAWLPRPAFSLTGLSLFFSLVPPGRSMEVTVPATLNVLNGDARLPCTFNSCYTVNHKQ
FSLNWTYQECNNCSEEMFLQFRMKIINLKLERFQDRVEFSGNPSKYDVSVMLRVNVQPEDEGI
YNCYIMNPPDRHRGHGKIHLQVLMEPPERDSTVAVIVGASVGGFLAVVILVLMVVKCVRRK
KEQKLSTDDLKTEEGKTDGEGNPDDGAK
```

Important features:

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 161-179

Immunoglobulin-like fold:

amino acids 83-127

N-glycosylation sites.

amino acids 42-45, 66-69 and 74-77

FIGURE 56

GTGTATATGCTGAAGTACATCCGTGCATTTTTAGCATCCAACCATCCTCCCTGTA
GTTCTGCCCTCAAATCACCTCTCCCTAGCCCACCCNACTAACATCTCAGTCTTGAA
AATGCACAGAGATGCCTGGCTACCTCGCCCTGCCTCAGCCTCACGGGGCTCAGTCTTTT
TCTCTTGGTGCACCAGGACGGAGCATGGAGGTCCACAGTACCTGNCCACCCTAACGTCC
TCAATGGCTCTGACGCCCGCCTGCCCTGCCCTCAACTCCTGCTACACAGTGAACCACAAAC
AGTTCTCCCTGAACTGGACTTACCAGGAGTGCAACAAC TGCTCTGAGGAGATGTTCCAG
TTCCGCATGAAGATCATTAAACCTGAAGCTGGAGCGGTTCAAGACCGCGTGGAGTTCTCAGG
GAACCCCAGCAAGTACGATGTGTCGGTGATGCTGAGAACGTGCAGCCGGAGGATGAGGGGA
TTTACAAC TGCTACATCATGAACCCCCC

FIGURE 57

TCACGGGGCTCATCTCTTTCTCTTGGTGCCACCAGGACGGAGCATGGAGGTNCACATA
CCTGCCACCCCTCAACGTCTCAATGGCTTGACGCCGCCTGCCCTGCACCTCAACTCCNG
CTACACAGTGAACCACAAACAGTTCTCCCTGAACCTGGATTTACCAGGAGTGCAACAACTGGC
TCTGAGGAGATGTTCCCTCCAGTTCCCGATGGAAGATCATTAAACCTGAAAGCTGGAAGCGG
TTTCAAGAACCGCGTGGAGTTCTCAGGAACCCAGCAAGTACGATGTGTCGGTGATGC
TGAGAAACGTGCAGCCGGAGGATGAGGGGATTACAACTGCTACATCATGAACCCCCC

FIGURE 58

TGCGGCGACCGTCGTACACCATGGGCCTCACCTCCGCCCTACCGTGTGGGCTGCTCCGGATGGCCTCCTGT
TCCTTGTGCTGCTAATGCTGCTCGCGAACCCAGCGCTCCGGCGACGTACCCCCCAGTGGTGTGCACTACCTCTGCTCCAAGAAGA
CCGAAAGCTACTTCACAATCTGGCTGAACCTGGAACCTGCTGCTGCCTGTCACTCATTGACTGCTGGATTGACAATA
TCAGGCTGGTTACAACAAAACATCCAGGGCCACCCAGTTCTGATGGTGTGGATGTACGTGTCCTGGCTTGGCTTGG
GGAAGACCTCTCACTGGAGTTCTGGACCCCAGCAAAAGCAGCGTGGGTCCCTATTCCACACCATGGTGGAGA
GCCCTGTGGCTGGGCTACACACGGGTGAGGATGTCCGAGGGCTCCCTATGACTGGGCCAGCCCCAAATG
AAAACGGGCCTACTTCTGGCCCTCCGCGAGATGATCGAGGAGATGTACAGCTGTATGGGGGCCCTGGTGC
TGGTTGCCACAGTATGGCAACATGTACACGCTCTACTTCTGCAAGCGGCAGCCAGGGCTGGAAGGACAAGT
ATATCCGGGCTTCGTGTCAGTGGTGCCTGGGGGGCTGGCCAAGACCCCTGGCGTCTGGCTTCAGGAG
ACAACAACCGGATCCCACTGGCCCTGAAGATCCGGAGCAGCAGCGGTCACTGTCTCCACCAGCTGGC
TGCTGCCCTACAACATGGCACCTGAGAAGGTGTCAGACACCCACAATCAACTACACACTGCGGG
ACTACCGCAAGTTCCAGGACATCGGCTTGAGATGGCTGCTCATGCGCAGGACACAGAAGGGCTGGTGG
AAGCCACGATGCCACCTGGCGTGCAGCTGCCTCATGGTACTGGCTCCACACAGAACTCCTTCTACT
ATGAGAGCTCCCTGACCGTGAACCTAAATCTGCTTTGGTGAAGGGATGGTACTGTGAAGAGTGGCC
TGCACTGGCAGGGCTGGCAGAGCCGAGGGACCAAGTGTGCTGCAAGGAGCTGCCAGGCAGGACATCG
AGATGCTGGCCAACGCCACCCCTGGCTATCTGAAACGTTGCTCTGGGGCCCTTGACTCCTGTGCCACAGGA
CTCCTGTGGCTCGGCCGTGGACCTGCTGTTGGCCTCTGGGGCTGTATGGCCACGGCTTTGCAAAGTTGTGA
CTCACCAATTCAAGGCCCGAGCTGGACTGTGAAGCATCTGCATGGGAAGTGCTGTTGTTATCCTTCTCT
GTGGCAGTGAGAAGAAGAAATGAGAGTCTAGACTCAAGGGACACTGGATGGCAAGAATGCTGCTGATGGTGG
ACTGCTGTGACCTTAGGACTGGCTCCACAGGGTGGACTGGCTGGCCCTGGTCCAGGCTCCCTGCTGGGGCCATG
TGTCCCCCTATTCTGTGGCTTTCTGCTACTGGCCCTGGCCCCGCAGCCTCCTATGAGGGATGTT
ACTGGGCTGTGGCTGTACCCAGAGGTCCAGGGATCGGCTCCTGGCCCTGGGTGACCCCTCCCACACACCA
GCCACAGATAGGCCTGCCACTGGCATGGTAGCTAGAGCTGCTGGCTTCCCTGGCTTAGCTGGTGGCCAGCC
TGACTGGCTCTGGCGAGCCTAGTAGCTCCTGCAGGCAGGGCAGTTGTCGTTCTCGTGGTCCCAGGC
CCTGGGACATCTCACTCCACTCCTACCTCCCTTACCAACAGGAGCATTCAAGCTCTGGATTGGGCAGCAGATGT
CCCCCAGTCCCGCAGGCTGTGTTCCAGGGCCCTGATTCCTGGATGTGCTATTGGCCCAGGACTGAAGCTGC
CTCCCTCACCTGGACTGTGGTCCAAGGATGAGAGCAGGGTGGAGCCATGGCCTCTGGGAACCTATGGA
GAAAGGGAAATCCAAGGAAGCAGCCAAGGCTGCTCGCAGCTCCCTGAGCTGCAACCTCTGCTAACCCACCATCA
CACTGCCACCCCTGCCCTAGGGTCTCACTAGTACCAAGTGGTCAGCACAGGGCTGAGGATGGGCTCCTATCCAC
CCTGGCCAGCACCAGCTTAGTGTGGACTAGCCCAGAAACTGAATGGGACCCCTGAGAGAGGCAGGGTCCCC
TGAGGGCCCCCTAGGGCTTCTGCTGCCAGGGTGTCCATGGATCTCCCTGTGGCAGCAGGCATGGAGAGT
CAGGGCTGCCTCATGGCAGTAGGCTCTAAGTGGGTGACTGGCACAGGCCAGAAAAGGGTACAGCCTCTAGGT
GGGGTCCCCAAAGACGCCCTCAGGCTGGACTGAGCTGCTCTCCACAGGGTTCTGTCAGCTGGATTTCTCTG
TTGCATACATGCCATCTGTCCTGGCTTCTGAGTGGCCCCACATGGGCTCTGAGCAGGCTGTATCTG
GATTCTGGCAATAAAAGTACTCTGGATGCTGTAaaaaaaaaaaaaaaaaaaaaaaa

FIGURE 59

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44189
><subunit 1 of 1, 412 aa, 1 stop
><MW: 46658, pI: 6.65, NX(S/T): 4
MGLHLRPYRVGLLPDGLLFLLLLLMLLADPALPAGRHPPVVLVPGDLGNQLEAKLDKPTVVH
YLCSSKKTESYFTIWLNLLELLLPVIIDCWIDNIRLVYNKTSRATQFPDGVDRVPGFGKTFSL
EFLDPSKSSVGSYFHTMVESLGVWGYTRGEDVRGAPYDWRRAPNENGPYFLALREMIEEMYQ
LYGGPVVLVAHSMGNMYTLYFLQRQPQAWKDKYIRAFVSLGAPWGGVAKTLRVLASGDNNRI
PVIGPLKIREQQRSAVSTSLLPYNTWSPEKVFVQTPTINYTLRDYRKFFQDIGFEDGWLM
RQDTEGLVEATMPPGVQLHCLYGTGVPTPDSFYYESFPDRDPKICFGDGDGTVNLKSALQCQ
AWQSRQEHQVLLQELPGSEHIEMLANATTLAYLKRVLLGP
```

Important features:

Signal peptide:

amino acids 1-28

Potential lipid substrate binding site:

amino acids 147-164

N-glycosylation sites.

amino acids 99-102, 273-276, 289-292 and 398-401

Lipases, serine proteins

amino acids 189-201

Beta-transducin family Trp-Asp repeat

amino acids 353-365

FIGURE 60

CGGACGCGTGGCGGACGCGTGGGGCGGCAGCGCGCGACGGCGAC **ATGGAGAGCGGG**
GCCTACGGCGGCCAAGGGGGCGGCTCCCTGACCTGCGCGCTTCCTGACGCAGCCGCA
GGTGGTGGCGCGCCGTGTGCTTGGTCTCGCCTGATCGTGTCTCCTGCATCTATGGT
AGGGCTACAGCAATGCCACGAGTCTAACGAGATGTACTGCGTGTCAACCGAACGAGGAT
GCCTGCCGCTATGGCAGTGCATCGGGGTGCTGGCCTCCTGGCCTCGGCCTTCTTCTGGT
GGTCACGCGTATTCCCCAGATCAGCAACGCCACTGACCGCAAGTACCTGGTCATTGGT
ACCTGCTCTCTCAGCTCTGGACCTTCCTGTGGTTGTGGTTCTGCTTCCTCACCAAC
CAGTGGGAGTCACCAACCGAAGGACGTGCTGGTGGGGCGACTCTGTGAGGGCAGCCAT
CACCTCAGCTTCTTCCATCTTCTCCTGGGGTGTGCTGGCCTCCCTGGCCTACCAGCGCT
ACAAGGCTGGCGTGGACGACTTCATCCAGAATTACGTTGACCCCACCTCCGGACCCAAACACT
GCCTACGCCCTACCCAGGTGCATCTGTGGACAACACTACCAACAGCCACCCCTCACCCAGAA
CGCGGAGACCACCGAGGGCTACCGCCGCCCCCTGTGTACT**TGAGTGGCGTTAGCGTGGAA**
GGGGGACAGAGAGGGCCCTCCCTCTGCCCTGGACTTCCCCTCAGCCTCTGGAACGTGCA
GCCCTCTTTCACCTGTCCATCCTGTGCAGCTGACACACAGCTAAGGAGCCTCATAGCC
TGGCGGGGGCTGGCAGAGCCACACCCAAAGTGCCTGTGCCAGAGGGCTTCAGTCAGCCGCT
CACTCCTCAGGGCACTTTAGGAAAGGGTTTAGCTAGTGTGTTCTCGCTTTAATGA
CCTCAGCCCCGCCTGCAGTGGCTAGAACGCCAGCAGGTGCCATGTGCTACTGACAAGTGCCT
CAGCTCCCCCGGCCGGTCAGGCCGTGGAGCCGCTATTATCTGCGTTCTGCCAAAG
ACTCGTGGGGCCATCACACCTGCCCTGTGCAGCGGAGCCGACCAGGCTTGTGCTCTCA
CTCAGGTTGCTTCCCTGTGCCACTGCTGTATGATCTGGGGCCACCACCCCTGTGCCGGT
GCCCTCTGGCTGCCTCCCGTGGTGTGAGGGCGGGCTGGTGCATGGCACTCCTCCTG
CTCCCACCCCTGGCAGCAGGGAAAGGGCTTGCCTGACAACACCCAGCTTATGAAATATTC
TGCAGTTGTTACTTAGGAAGGCCTGGGAGGGCAGGGGTGCCCATGGCTCCAGACTCTGTC
TGTGCCGAGTGTATTATAAAATCGTGGGGAGATGCCCGGCCTGGGATGCTGTTGGAGACG
GAATAAAATGTTTCTCATTCAAAG

FIGURE 61

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48304
<subunit 1 of 1, 224 aa, 1 stop
<MW: 24810, pI: 4.75, NX(S/T): 1
MESGAYGAAKAGGSFDLRRFLTQPQVVARAVCLVFALIVFSCIYGEGYSNAHESKQMYCVFN
RNEDACRYGSAIGVLAFLASAFFLVDAYFPQISNATDRKYLVIGDLLFSALWTFLWFVGFC
FLTNQWAVTNPKDVLVGADSVRAAITFSFFSIFSWGVLASLAYQRYKAGVDDFIQNYVDPTP
DPNTAYASYPGASVDNYQQPPFTQNAETTEGYQPPPVY
```

Important features:

Type II Transmembrane domain:

amino acids 1-45

Other transmembrane domains:

amino acids 74-90, 108-126 and 145-161

N-glycosylation site.

amino acids 97-100

FIGURE 62

GAGCCACCTACCCCTGCTCCGAGGCCAGGCCAGGGCTGCAGGCCATCGGCCAGAGGGTATCAGTGAGCAGAAGGATG
CCCGTGGCCGAGGCCCCCAAGGGCTGGCGGGCAGGGGACGGGTATGGCGAGGAAGCGGAGCCAGAGGG
ATGTTCAAGGCCTGTGAGGACTCCAAGAGAAAAGGCCGGGCTACCTCCGCCCTGGTGCCTGCTGCTG
GCCCTGCTCGTCTGGCTCGGGGGGTGCTACTCTGGTATTCTCTAGGGTACAAGGCCAGGGTATGGTCAGC
CAGGTGTACTCAGGCAGTCTCGCTGTACTCAATGCCACTCTCCAGGATCTTACCCGCCGGAACTCTAGTGCC
TTCCGCAGTGAACCGCAAAGCCCAGAAGATGCTCAAGGAGCTCATCACAGCACCCGCTGGAACTTACTAC
AACTCCAGCTCCGTCTATTCCCTTGGGGAGGGACCCCTCACCTGCTTCTGGTTCATCTCAAATCCCCGAG
CACCGCCGGTGTGAGCCCCGAGGTGGTGCAGGCAGTCTGGTGGAGGAGCTGCTGTCCACAGTCACAGC
TCGGCTGCCGCTCCCTACAGGGCCAGTACGAAGTGGACCCCGAGGGCTAGTGTACCTGGAAAGCCAGTGTGAAA
GACATAGCTGATTGAATTCCACGCTGGGTTTACCGCTACAGCTACGTGGCCAGGGCCAGGTCCCTCCGGCTG
AAGGGGCTGACCACCTGGCCTCAGCTGCCTGTGGCACCTGCAGGGCCCAAGGACCTCATGCTCAAACCTCCGG
CTGGAGTGGACGCTGGCAGAGTGGCAGGACTGGCATGTATGACGTGGCCGGGGCCCTGGAGAAGAGGCTC
ATCACCTCGGTGTACGGCTGCAGCCGCCAGGAGGCCGTGGTGGAGGTTCTGGCTCGGGGCCATCATGGCGTC
GTCTGGAAGAAAGGGCCTGCAAGCTACTACGACCCCTTCGTGCTCCGTGCAAGCCGGTGGTCTTCAGGCCTGT
GAAGTGAACCTGACGCTGGACAAACAGGCTGACTCCAGGGCTCCAGCACCCCGTACTTCCCAGCTACTAC
TCGCCCAAACCCACTGCTCTGGCACCTCACGGTGGCCCTCTGGACTACGGCTTGGCCCTCTGGTTGATGCC
TATGCACTGAGGAGGCCAGAAGTATGATTGGCGTGCACCCAGGGCCAGTGACGATCCAGAACAGGAGGCTGTG
GGCTGCGCATCTGCAAGCCCTACGCCAGGAGGATCCCCGTGGCCACGGCCGGGATCACCATCAACTTACCC
TCCCAGATCTCCCTACCGGGCCGGTGTGGGGTGACTATGGCTTGACAAACAGCTGGACCCCTGGCCCTGGA
GAGTTCCCTGTTCTGTAATGGACTCTGTGTCCTGCTGTGATGGGTCAAGGACTGCCAACGGCCTGGAT
GAGAGAAACTGCGTTGCAGAGCCACATTCCAGTGCAAAGAGGACAGCACATGCATCTCACTGCCAACGGTCTGT
GATGGGAGGCTGATTGCTCAACGGCAGCGATGAAGAGCAGTGCCAGGAAGGGTGGCCATGTGGGACATTACCC
TTCCAGTGTGAGGACGGAGCTGCGTGAAGAAGCCAAACCCGCAAGTGTGATGGGCGGGCCACTGCAGGGACGGC
TCGGATGAGGAGCACTGTGACTGTGGCTCCAGGGCCCTCCAGCCGCAATTGGTGGAGGCTGTGCTCCGAG
GGTAGTGGCCATGGCAGGCCAGCTCCAGGTTGGTGCACACATCTGTGGGGGGCCCTCATGCTGACCGC
TGGGTGATAACAGCTGCCACTGCTTCAGGAGGACAGCATGGCTCCACGGTGTGTTGGACCGTGTGCTGGG
AAGGTGTGGCAGAACTCGCGTGGCTGGAGAGGTGCTTCAAGGTGAGCCGCTGCTCCCTGACCCGTACAC
GAAGAGGACAGCCATGACTACGACGTGGCGTGCAGCTCGACCACCCGGTGGTGCCTCGCCGCGTGC
CCCGTCTGCTGCCCGCGCTCCACTCTCGAGGCCGCTGCACTGCTGGATTACGGCTGGGGCGCTTG
CGCGAGGGCGGCCATCAGCAACGCTCTGCAGAAAGTGGATGTGCACTGGTGTGCTCCACAGGACCTGTGAGCAG
GCCTATCGCTACCAGGTGACGCCACGCATGCTGTGCGGCTACCGCAAGGGCAAGAAGGATGCCGTGAGGT
GACTCAGGTGGTCCGCTGGTGTGCAAGGCACTCAGTGGCGCTGGTCTGGGGCTGGTCAAGCTGGGGCTG
GGCTGTGGCCGCTAACTACTCTGGCGTCTACACCCGCACTCACAGGTGTGATCAGCTGGATCCAGCAAGTGGT
ACCTGAGGAACTGCCCTGCAAAGCAGGGCCACCTCTGGACTCAGAGAGGCCAGGGCAACTGCCAACGAG
GGGACAAGTATTCTGGGGGGGGGGAGAGGAGCAGGGCTGTGGTGGCAGGAGGTGGCATCTGTCTCGTCC
CTGATGTCTGCTCCAGTGTGGCAGGAGGATGGAGAAGTGGCAGCAGCTGGGGGTCAAGACGTCCCTGAGGACC
CAGGCCACACCCAGCCCTCTGCCCTCCAAATTCTCTCCTCCCTCCCTCCACTGCTGCCATAATGCAAG
GCAGTGGCTCAGCAGCAAGAATGCTGGTTTACATCCGAGGAGGTGTGAGGTGCGCCCCACTCTGTACAGAGG
CTGTTGGCAGCCTTGCTCCAGAGGAGCAGATTCCAGCTCGGAAGCCCTGGTCAACTTGGGATCTGGGAAT
GGAAGGTGCTCCCATCGGAGGGACCCCTCAGAGCCCTGGAGAGACTGCCAGGTGGCCTGCCCCACTGTAAGCCAA
AAGGTGGGAAAGTCTGACTCCAGGGTCTGGCCCCACCCCTGCCACCTGGCCTCACAGCCCAGACCC
CACTGGGAGGTGAGCTCAGCTGCCCTTGGAAATAAGCTGCCATGATCAAAAAAAAAAAAAAA

FIGURE 63

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49152
><subunit 1 of 1, 802 aa, 1 stop
><MW: 88846, pI: 6.41, NX(S/T): 7
MPVAEAPQVAGGQGDGGDGEAEPEGMFKACEDSKRKARGYLRLVPLFVLLALLVLASAGVL
LWYFLGYKAEVMVSQVYSGSLRVLNRHFSQDLTRRESSAFRSETAKAQKMLKELITTRLGT
YYNSSSVYSFGEGLTCFFWFILQIPEHRRRLMLSPEVVQALLVEELLSTVNSSAAVPYRAEY
EVDPEGLVILEASVKDIAALNSTLGCYRYSYVGQGQVLRLKGPDHLASSCLWHLQGPKDLML
KLRLEWTLAECRDRRLAMYDVAGPLEKRLITSVYGCSRQEPMVEVLAGAIMAVVWKGLHSY
YDPFVLSVQPVVFQACEVNLTLDNRLDSQGVLSPTYFPSYYSPQTHCSWHLTVPSLDYGLAL
WFDAYALRRQKYDLPCTQGQWTIQNRRLCGLRILQPYAERIPVVATAGITINFTSQISLTGP
GVRVHYGLYNQSDPCPGEFLCSVNGLCVPACDGVKDCPNGLDERNCVCRATFQCKEDSTCIS
LPKVCDGQPDCLNGSDEEQCQEGVPCGTFTQCEDRSCVKKPNPQCDGRPDCRDGSDEEHCD
CGLQGPSSRIVGGAVSSEGEWPWQASLQVRGRHICGGALIADRWVITAACFQEDSMASTVL
WTVFLGKVWQNSRWPGEVSKVSRLLLHPYHEEDSHDYDVALLQLDHPVVRSAAVRPVCLPA
RSHFFEPGLHCWITGWGALREGGPISNALQKVDVQLIPQDLCSEAYRYQVTPRMLCAGYRKG
KKDACQGDSGGPLVCKALSGRWFLAGLVSWGLGCRPNYFGVYTRITGVIWSIQQVVT
```

Important features:

Type II transmembrane domain:

amino acids 46-67

Serine proteases, trypsin family, histidine active site.

amino acids 604-609

N-glycosylation sites.

amino acids 127-130, 175-178, 207-210, 329-332, 424-427, 444-447
and 509-512

Kringle domains.

amino acids 746-758 and 592-609

Homologous region to Kallikrein Light Chain:

amino acids 568-779

Homologous region to Low-density lipoprotein receptor:

amino acids 451-567

FIGURE 64

GCACCCAGGGCAGTGGACGATCCAGAACAGGAGGCTGTGGCTTGCATCCTGCAGCCC
TACGCCGAGAGGATCCCCGTGGTGGCCACGGCCGGGATCACCATCAACTCACCTCCAGAT
CTCCCTCACCGGGCCGGTGTGCAGGGTGCACATGGCTGTACAACCAGTCGGACCCCTGCC
CTGGAGAGTTCTCTGTTCTGTGAATGGACTCTGTGTCCCTGCCTGTGATGGGTCAAGGAC
TGCCCCAACGGCCTGGATGAGAGAAACTGCCTTGAGGCCACATTCCAGTGCAAAGAGGA
CAGCACATGCATCTCACTGCCAACGGTCTGTGATGGCAGGCCCTGATTGTCTAACGGCAGCG
ATGAAGAGCAGTGCCAGGAAGGGTGCCATGTGGACATTCACCTCAGTGTGAGGACCGG
AGCTGCGTGAAGAACGCCAACCGCAGTGTGATGGCGGCCACTGCAGGGACGGCTCGGA
TGAGGAGCACTGTGACTGTGGCCTCCAGGGCCCTCCAGGCCGATTGTTGGTGGAGCTGTGT
CCTCCGAGGGTGAGTGGCCATGGCAGGCCAGCCTCCAGGTTGGGTCGACACATCTGTGGG
GGGCCCTCATCGCTGACCGCTGGTGATAAACAGCTGCCACTGCTTCAGGAGGACAGCAT
GGCCTCCACGGTGTGGACCGTGTTCCTGGCAAGGTGTGGCAGAACTCGCGCTGGCCTG
GAGAGGTGTCTTCAAGGTGAGCCGCTGCTGCAGCTCGACCACCCGGTGGTGCCTGGCCGGTGC
GACTACGACGTGGCGCTGCTGCAGCTCGACCACCCGGTGGTGCCTGGCCGGTGC
CGTCTGCCTGCCCGCGCTCCACTTCTCGAGCCGCTGCAGTGTGGATTACGGCT
GGGCGCCTTGCAGGGCGGCCATCAGCAACGCTCTGCAGAAAGTGGATGTGCAGTTG
ATCCCACAGGACCTGTGCAGCGAGGCCTATCGTACCAAGGTGACGCCACGCATGCTGTGTGC
CGGCTACCGCAAGGGCAAGAAGGATGCCTGTCAAGGTGACTCAGGTGGTCCGCTGGTGTGCA
AGGCACTCAGTGGCGCTGGTCTGGCGGGCTGGTCAGCTGGGCTGGCTGTGGCG
CCTAACTACTCGCGTCTACACCCGATCACAGGTGTGATCAGCTGGATCCAGCAAGTGGT
GACCTGAGGAAGTGCCCTGCAAAGCAGGGCCACCTCCTGGACTCAGAGAGCCAGGGC
AACTGCCAAGCAGGGGACAAGTAT

FIGURE 65

GGACGAGGGCAGATCTCGTTCTGGGGCAAGCCGTTGACACTCGCTCCCTGCCACCGCCCCGGG
CTCCGTGCCGCCAAGTTTCATTTCCACCTCTGCCTCCAGTCCCCCAGCCCCCTGGCCG
AGAGAAGGGTCTTACCGGCCGGATTGCTGGAAACACCAAGAGGTGGTTTTGTTTTAAA
ACTTCTGTTCTTGGGAGGGGGTGTGGCGGGCAGGATGAGCAACTCCGTTCTGCTCTG
TTTCTGGAGCCTCTGCTATTGCTTGCTGCCGGAGCCCCGTACCTTTGGTCCAGAGGGAC
GGCTGGAAGATAAGCTCCACAAACCAAAGCTACACAGACTGAGGTCAAACCATCTGTGAGG
TTAACCTCCGACCTCCAAGGACCCAGAGCATGAAGGATGCTACCTCTCCGTCGGCACAG
CCAGCCCTAGAAGACTGCAGTTCAACATGACAGCTAAACCTTTCATCATTACGGAT
GGACGATGAGCGGTATTTGAAAAGCTGGCTGCACAAACTCGTGTCAAGCCCTGCACACAAGA
GAGAAAGACGCCAATGTAGTTGTGGTTGACTGGCTCCCCCTGGCCCACCAGCTTACACGGA
TGC GGTCATAATAACCAGGGTGGTGGACACAGCATTGCCAGGATGCTGACTGGCTGCAGG
AGAAGGACGATTTCTCGGAAATGTCCACTTGATCGGCTACAGCCTCGGAGCGCACGTG
GCCGGGTATGCAGGCAACTCGTGAAGGAACCGTGGGCCAATCACAGGTTGGATCCTGC
CGGGCCCATGTTGAAGGGGCCGACATCCACAAGAGGCTCTCCGGACGATGCAGATTTG
TGGATGTCCCTCACACCTACACGCGTTCTCGGCTTGAGCATTGGTATTAGATGCCTGTG
GCCACATTGACATCTACCCAAATGGGGTGACTTCCAGCCAGGCTGTGGACTCAACGATGT
CTTGGGATCAATTGCATATGGAACAATCACAGAGGTGGTAAAATGTGAGCATGAGCGAGCCG
TCCACCTTTGTTGACTCTCGGTGAATCAGGACAAGCCGAGTTGCCTCAGTGCACT
GACTCCAATCGCTCAAAAGGGATCTGCTGAGCTGCCAAGAACCGTTGTAATAGCAT
TGGCTACAATGCCAAGAAAATGAGGAACAAGAGGAACAGCAAAATGTACCTAAAAACCCGGG
CAGGCATGCCCTTCAGAGGTAACCTCAGTCCCTGGAGTGTCCCTTGAGGAAGGCCCTTAATA
CCTCCTTCTTAATACCATGCTGCAGAGCAGGGCACATCCTAGCCCAGGAGAAGTGGCCAGCA
CAATCCAATCAAATCGTTGCAAATCAGATTACACTGTGCATGTCCTAGGAAAGGGAATCTT
ACAAAATAAACAGTGTGGACCCCTAATAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAA

FIGURE 66

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49646
><subunit 1 of 1, 354 aa, 1 stop
><MW: 39362, pI: 8.35, NX(S/T): 2
MSNSVPLLCFWSLCYCFAAGSPVPGPEGRLEDKLHKPKATQTEVKPSVRFLRTSKDPEHE
GCYLSVGHSQPLEDCSFNMTAKTFFIIHGWTMSGIFENWLHKLVSLHTREKDANVVVDWL
PLAHQLYTDAVNNTRVVGHSIARMLDWLQEKKDDFSLGNVHLIGYSLGAHVAGYAGNFVKGT
GRITGLDPAGPMFEGADIHKRLSPDDADFVDVLHTYTRSFGLSIGIQMPVGHIDIPNGGDF
QPGCGLNDVLGSIAYGTITEVKCEHERAVHLFVDSLNVQDKPSFAFQCTDSNRFKKGICLS
CRKNRCNSIGYNAKKMRNKRNSKMYLKTRAGMPFRGNLQSLECP
```

Important features:

Signal peptide:

amino acids 1-16

Lipases, serine active site.

amino acids 163-172

N-glycosylation sites.

amino acids 80-83 and 136-139

FIGURE 67

CGGACGCGTGGCGGACGCGTGGCCTGGCAAGGGCGGGCGCCGGCGAGCCACCTTTCCTCCCCCGCT
TCCCTGCGCTCCGCTGGCTGGACGCGCTGGAGGAGTGGAGCAGCACCCGGCCGCCCTGGGGCTGACAGT
CGGCAAAGTTGGCCCGAAGAGGAAGTGGTCTCAAACCCGGCAGGTGGGACCCAGGGCAGACCAGGGCGCTCG
CTGCCTGCGGGCGGGCTGTAGGGAGGGCGCCCGAGTGCAGGAGACCCGGGCTTCAGGAGGCCCGGGAG
AGAAGAGTGCAGGGCGGAGGGAAAACAACCTCAAAGTTGGGAAAGGCACCGCCCTACTCCCGGGCTGCCG
CCGCCTCCCCGCCCGAGCCCTGGCATCCAGAGTACGGGTGAGCCCGGGCATGGAGCCCCCTGGGAGGCG
CACCAGGGAGCCTGGCGCCGGGCTCCGCCGAGCCCATGGTAGACCACAGAACGCTCCGGGACCCCTCCG
GCACCTCTGACAGCCAGGATGCTGTTGGCACCCCTCCTCCCTCTGGAGGCCTGGCCCATCCAG
ACCGGATTATTTTCAAATCATGCTTGTGAGGACCCCCAGCAGTGCTCTAGAAGTGCAGGGCACCTACAGA
GGCCCTGGTCCGGGACAGCCGACCTCCCTGCCACTGCACCTGGCTCATCTGGCAGCAAGGAACAGACTG
TCACCATCAGGTTCCAGAAGCTACACCTGGCTGTGGCTAGAGCCTTAACCTACGCTCCCTCAGCCAC
TGATCTCCCTGTTGAGGACCTCCAGCCCTCTGCAGCTGCCGGGGCACGTCACCACACTACAGCTATG
CTGGGCCAGAGCACCCATGGGCCAGGGCTTCCTGCTCTACAGCCAAGATTGGCTGATGTGCCCTGCAGGAAG
AGTTTCAGTGCCTGAACCACCGCTGTGATCTGCTGCTGCCAGCGCTGTGATGGGTTGATGCCCTGTCAGGATGGCT
CTGATGAGCAGGGTGCAGCTCACCCCTCCCTGCCGATGCCAACAGACCCGTCCTCCCTGCCCTGCAATG
TCACCTTGGAGGACTTATGGGTCTCTCTCTGGATATCACACACTAGCCTCAGTCTCCACCCCTGAG
CCTGCCATTGGCTGGCACCCATGATGGCCGGCTGGCGCTACAGGCCCTGGACACTTGGGCTTGG
GAGATGCAGTGCATGTGATGACGGCCCTGGGCCCCCTGAGAGCTCCGACTACTGCGTAGTCTCACCCACTTCA
GCAATGGCAAGGCTGTCACTGTGGAGACACTGTCTGGCAGGCTGTTGTCCTACCAACAGTTGCTTGGAGCA
ATGGTCTGCTTCAATGCCACCTACCATGTGCGGGCTATTGCTGCTTGGACAGACCCCTGGCTTAGGCT
CTGGCCTGGAGCTGGCAAGGCCTAGGTGAGCCTGCTACAGTGAGGACAGCGCTGTGACGGCTATGGGACT
GTGCTACGGCACAGATGAGGAGACTGCCAGGCTGCCACCTGGACACTTCCCTGTCGGGCTGCTGGCACCT
CTGGTGCACAGCCTGCTACCTGCCGCTGCAACTACCAGACTTCTGTGCTGATGGAGCAGATGAGA
GACGCTGTGGCATTGCCAGCCTGGCAATTCCGATGCCGGACGAGAAGTGCCTGATGAGACGTGGGTGCG
ATGGCAGCCAGACTGTGCGGACGGCAGTGATGAGTGGACTGCTCTATGTTCTGCCCGCAAGGTATTACAG
CTGCAGTCATTGGCAGCCTAGTGTGCGGCCCTGCTCTGGTATGCCCTGGCTGACACTCTATGCCA
TTCGCACCCAGGAGTACAGCATCTTGCCCCCTCCCGATGGAGGCTGAGATTGTGCGAGCAGGCA
CTTCCCTACGGGAGCTCATTGCCAGGGTGCATCCACCTGTTAGAAAGACTTCCCTACAGAGAACTTAATGATA
ACTCAGTGTGGCAACCTGCGTCTCTGCTACAGATCTACGCCAGGATATGACTCCAGGAGGTGGCCAGGT
CCCGCCGTGTCAGCGGGCCGCTTGATGCGACGCCCTGGTACGCCGCTCCGCCGCTGGGCTTGCTCCCTCGAA
CCAACACCCCGGCTCGGGCTCTGAGGCCAGATCCAGGTACACCTCTGCTGCTCCCTTGAGGCCCTAGATG
GTGGCACAGGTCCAGCCGTGAGGGCGGGCAGTGGGTGGCAAGATGGGAGCAGGCCACCCACTGCCCATCA
AGGCTCCCTCCCATCTGCTAGCACGCTCAGCCCCACTACTGCTCCCTGAAGGCCAGGGCCACTGCCCTCAC
TGCCCTAGAGCCATCACTATTGCTGGAGTGGTGCAGGCCCTGCGAGGCCGCTGTGCCCAGGCTGGGCCCC
CAGGACCAACCCGGAGCCCCCTGGACCCACACAGCAGTCTGCCCTGGAAGATGAGGAGCAGTGTGCTACTGG
TGCCACTGGCTGAGGGGGGTGTTAGCTGAGGAGAGGATGAGCCACTGCTTACCTGAGGGACCTGGGG
CTCTACTGAGGCCCTCTCCCTGGGGCTACTCATAGTGGCACAAACCTTATAGAGTGGCTAGCCTCCCCCTCC
ACCAACTCCTCCCTGCTGGATTTCAGGGACTTGGTGGGCCCTCCGTTGACCTATGTTAGCTGCTATAAAGT
TAAGTGTCCCTCAGGCAGGGAGAGGGCTCACAGAGTCTCTGCTACGTGCCATGGCAGACACCCAGTCCCT
TCACCAACCACTGCTCCCCAGCCACCACTTGGGTGGCTGTTTAAAGTAAAGTCTTAGAGGATCATA
GGTCTGGACACTCCATCTGCCAACCTTACCCAAAAGTGGCTTAAGCAGCGGAATGCCAATTAACTAGAGA
CCCTCCAGCCCCAAGGGAGGATTGGGAGAACCTGAGGTTTGCCATCCACAATCCCTACAGGGCCTGG
CTCACAAAAAGAGTGCACAAATGCTTCTATTCCATAGCTACGGCATTGCTCAGTAAGTTGAGGTAAAAATAAA
GGAATCATACTC

FIGURE 68

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49631
<subunit 1 of 1, 713 aa, 1 stop
<MW: 76193, pI: 5.42, NX(S/T): 4
MLLATLLLLLGGALAHPDRIIFPNHACEDPPAVLLEVQGTLQRPLVRDSRTSPANCTWLIL
GSKEQTVTIRFQKLHLACSERLTLRSPLQPLISLCEAPPSPQLPGGNVTITYSAGARAP
MGQGFLLSYSQDWLMCLQEEFQCLNHRCVSAVQRCDGVDACGDGSDEAGCSSDPFPGLTPRP
VPSLPCNVTLLEDFYGVFSSPGYTHLASVSHPQSCHWLLDPHDGRRLAVERFTALDLGFGDAVH
VYDGPGPPESSRLLRSLTHFSNGKAVTVETLSGQAVVSYHTVAWSNGRGFNATYHVRGYCLP
WDRPCGLGSGLGAGEGLGERCYSEAQRCDGSWDCADGTDEEDCPGCPPGHFPCGAAGTSGAT
ACYLPADRCNYQTFCADGADERRCRHCQPGNFRCRDEKCVYETWVCDGQPDCAADGSDEWDCS
YVLPRKVITAAVIGSLVCGLLLVIALGCTCKLYAIRTQEYSIFAPLSRMEAEIVQQQAPPSY
GQLIAQGAIPPVEDPTENPNDNSVLGNLRSLLQILRQDMTPGGGPGARRRQRGRLMRRLVR
RLRRWGLLPRNTPARASEARSQVTPSAAPLEALDGGTGPARSEGGAVGGQDGEQAPPLPIKA
PLPSASTSPAPTVPEAPGPLPSLPLEPSLLSGVVQALRGRLLPSLGPPGTRSPPGPHTAV
LALEDEDVLLVPLAEPGVWVAEAEDEPLLT
```

Important features:

Signal peptide:

amino acids 1-16

Transmembrane domain:

amino acids 442-462

LDL-receptor class A (LDLRA) domain proteins

amino acids 411-431, 152-171, 331-350 and 374-393

FIGURE 69

CGAGCTGGCGAGAAGTAGGGAGGGCGGTGCTCCGCCGCGTGGCGTTGCTATCGCTTCG
CAGAACCTACTCAGGCAGCCAGCTGAGAAGAGTTGAGGGAAAGTGTGCTGCTGGGTCTGCA
GACGCGATGGATAACGTGCAGCGAAAATAACATCGCCCTTCTGCTTCAGTGTGAAAGG
CCACGTGAAGATGCTGCGGCTGGCACTAACTGTGACATCTATGACCTTTTATCATCGCAC
AAGCCCCCTGAACCATATATTGTTATCACTGGATTGAAGTCACCGTTATCTTATTTTCTATA
CTTTTATATGTACTCAGACTTGATCGATTAATGAAGTGGTTATTTGGCCTTGCTGATAT
TATCAACTCACTGGTAACAACAGTATTGATGCTCATCGTATCTGTGTTGGCACTGATAACCAAG
AAACCACAAACATTGACAGTTGGGGTGGAGGGGTGTTGCACTTGTGACAGCAGTATGCTGTCTT
GCCGACGGGGCCCTTATTACCGGAAGCTCTGTTCAATCCCAGCGGTCTTACCAAGAAAAAA
GCCTGTGCATGAAAAAAAAGAAGTTTGTAATTTATATTACTTTAGTTGATAACTAAGT
ATTAAACATATTCTGTATTCTCCAAAAAA

卷之三

FIGURE 70

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49645

><subunit 1 of 1, 152 aa, 1 stop

><MW: 17170, pI: 9.62, NX(S/T): 1

MDNVQPKIKHRPFCFSVKGHVKMLRLALTTSMTFFIIAQAPEPYIVITGFEVTVILFFILL
YVLRLDRLMKWLFWPLLDIINSLVTTVFMLIVSVLALIPETTTLTVGGGVFALVTAVCCLAD
GALIYRKLLFNPSGPYQKKPVHEKKEVL

Important features:

Potential type II transmembrane domain:

amino acids 26-42

Other potential transmembrane domain:

amino acids 44-65, 81-101 and 109-129

Leucine zipper pattern

amino acids 78-99 and 85-106

N-myristoylation site.

amino acids 110-115

Ribonucleotide reductase large subunit protein

amino acids 116-127

FIGURE 71

GGGCAGAGAAGTAGGGGAGGGCGTGTCCGCCGCGGTGGCGGTTGCTATCGTTGCAGAACCTACTCAGGCAGCCAGNTGAGAAAGAGTTGAGGGAAAGTGCTGCTGGTCTGCAGACGCGATGGATAACGTGCAGCCGAAAATAAAACATCGCCCCCTCTGCTTCAGTGTGAAAGGCCACGTGAAGATGCTGCGGCTGGCACTAACTGNGACATCTATGACCTTTTATNATCGCACAAGCCCCTGAACCATAATTGTTACTGGATTGAAGTCACCGTTATCTTATTTTCATACTTTATATGTAACTCAGACTTGATCGATTAATGAAGTGGTTATTTGGCCTTGCTGATATTATCAACTCACTGGTAACAACAGTATTGCTCATCGTATCTGTGTTGGCACTGATACCAGAAACCACAACATTGACAGTTGGTGGAGGGGTGTTGCACTGTGACAGCAGTATGCTGTNTTGGCGAC

卷之三

FIGURE 72

CAGCCCCGCGCGCCGGCCGAGTCGCTGAGCCGCGCTGCCGGACGGGACGGGACCGGCTAGG
CTGGGCGCGCCCCCGGGCCCGCCGTGGGATGGGCGCACTGGCCCGGGCGCTGCTGCTGC
CTCTGCTGGCCAGTGGCTCCTGCGCGCCGCCGGAGCTGGCCCCCGGCCCTCACGCTG
CCCCTCCGGGTGGCCGCGGCCACGAACCGCGTAGTTGCGCCCACCCCGGGACCCGGACCC
TGCCGAGCGCCACGCCGACGGCTTGGCGCTGCCCTGGAGCCTGCCCTGGCGTCCCCCGCG
GCGCCGCCAACCTCTGGCCATGGTAGACAAACCTGCAGGGGACTCTGGCCGGCTACTAC
CTGGAGATGCTGATCGGGACCCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAG
TAACTTGCCTGGCAGGAACCCCGACTCCTACATAGACACAGTACTTGACACAGAGAGGT
CTAGCACATACCGCTCCAAGGGCTTGACGTACAGTGAAGTACACACAAGGAAGCTGGACG
GGCTTCGTTGGGAAGACCTCGTCACCATCCCAAAGGCTTCAATACTTCTTGTCAA
CATTGCCACTATTTGAATCAGAGAATTCTTTGCCTGGATTAAATGGAATGGAATAC
TTGGCCTAGCTTATGCCACACTGCCAAGCCATCAAGTTCTGGAGACCTTCTCGACTCC
CTGGTACACAAGCAAACATCCCAAACGTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCC
CGTTGCTGGATCTGGGACCAAACGGAGGTAGTCTGTCTGGTGGATTGAACCAAGTTGT
ATAAAGGAGACATCTGGTATAACCCCTATTAGGAAGAGTGGTACTACCAGATAGAAATTCTG
AAATTGGAAATTGGAGGCCAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGC
CATCGTGGACAGTGGCACCACGCTGCGCCTGCCCAAGAGGTGTTGATCGGTGGTGG
AAGCTGTGGCCCGCGCATCTGATTCCAGAATTCTCTGATGGTTCTGGACTGGTCCCAG
CTGGCGTGCTGGACGAATTGGAAACACCTGGTCTTACTCCCTAAATCTCCATCTACCT
GAGAGACGAGAACTCCAGCAGGTATTCCGTATCACAATCTGCCTCAGCTTACATTAGC
CCATGATGGGGCCGGCCTGAATTATGAATGTTACCGATTGGCATTCCCCATCCACAAAT
GCGCTGGTATCGGTGCCACGGTATGGAGGGCTTCTACGTATCTCGACAGAGCCAGAA
GAGGGTGGGCTTCGCAAGCGAGCCCTGTGCAGAAATTGCAGGTGCTGCAGTGTCTGAAATT
CCGGGCCTTCTAACAGAGGATGTAGCCAGCAACTGTGTCCCCGCTCAGTCTTGAGCGAG
CCCATTGTGGATTGTGTCTATGCGCTCATGAGCGTCTGTGGAGCCATCCTCTGTCTT
AATCGTCCTGCTGCTGCCGTTCCGGTGTGCAGCGTCGCCCGTGACCGTCAAGGTCGTCA
ATGATGAGTCCTCTGGTCAGACATCGCTGGAAATGAATGCCAGGCTGACCTCAAGCAA
CCATGAACTCAGCTATTAGAAAATCACATTCCAGGGCAGCAGCCGGATCGATGGTGGCG
CTTCTCTGTGCCAACCGTCTTCAATCTGTTCTGCTCCAGATGCCTTAGATTAC
TGTCTTTGATTCTGATTTCAGCTTCAAAGCTTCAAATCCTCCCTACTTCCAAGAAAAATAATTAA
AAAAAAAACCTCATTCTAA

FIGURE 73

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45493
><subunit 1 of 1, 518 aa, 1 stop
><MW: 56180, pI: 5.08, NX(S/T): 2
MGALARALLPLLAQWLLRAAPELAPAPFTLPLRVAATNRVVAPTPGPGTAAERHADGLAL
ALEPALASPAGAANFLAMVDNLQGDSGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAGTPHS
YIDTYFDTERSSTYRSKGFDTVVKYTQGSWTGFVGEDLVTIPKGFNTSFLVNIATIFESEN
FLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCAGLPVAGSGTNGGS
LVLGGIEPSLYKBDIWTPIKEEWYYQIEILKLEIGGQSLNLCREYNADKAIVDSGTTLLR
LPQKVFDAVVEAVARASLIPEFSDGFWTGSQLACWTNSETPWSYFPKISIYLRDENSRSFR
ITILPQLYIQPMMGAGLNCEYRGISPSTNALVIGATVMEGFYVIFDRAQKRVGFAASPCA
EIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCGAILLVLIVLLLLPFRC
QRPRDPEVVNDESSLVRHRWK
```

Important features:

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 466-494

N-glycosylation sites.

amino acids 170-173 and 366-369

Leucine zipper pattern.

amino acids 10-31 and 197-118

Eukaryotic and viral aspartyl proteases

amino acids 109-118, 252-261 and 298-310

FIGURE 74

CGCCTCCGCCTCGGAGGCTGACGCGCCGGCGCCGTTCCAGGCCTGTGCAGGGCGGATCG
GCAGCCGCCTGGCGGCGATCCAGGGCGGTGCGGGGCCTGGCGGGAGCCGGGAGGCACGGCC
GGC**ATGG**AGGCCTGCTGCTGGCGGGGTTGCTGCTGGCGCTTACGTGCTTGTCTACTA
CAACCTGGTGAAGGCCCGCCGTGCGGCGGCATGGCAACCTGCGGGCCGACGGCGTGG
TCACGGCGCCAACAGCGGCATCGAAAGATGACGGCGCTGGAGCTGGCGCGCCGGGAGCG
CGCGTGGTGCTGGCCTGCCGAGCCAGGAGCGCGGGAGGCCTGCCTCGACCTCCGCA
GGAGAGTGGAAACAATGAGGTACATTCATGGCCTGGACTTGGCCAGTCTGGCCTCGGTG
GGGCCTTGCCACTGCCTTCTGAGCTCTGAGCCACGGTTGGACATCCTCATCCACAATGCC
GGTATCAGTTCTGTGGCGGACCGTGAGGCCTTAACCTGCTGCTTCGGGTGAACCATA
CGGTCCCTTCTGCTGACACATCTGCTGCTGCCTGCGCTGAAGGCATGTGCCCTAGCCGCG
TGGTGGTGGTAGCCTCAGCTGCCACTGTGGGACGTCTGACTTCAAACGCCCTGGACCGC
CCAGTGGTGGCTGGCGCAGGAGCTGCGGCATATGCTGACACTAACGCTGGCTAATGTACT
GTTTGCCCGGGAGCTGCCAACCAGCTTGAGGCCACTGGCGTACCTGCTATGCAGCCCACC
CAGGGCCTGTGAACCTGGAGCTGTTCTGCCATGTTCTGGATGGCTGCCACTTTG
CGCCCATTGGCTGGCTGGTGCTCCGGCACCAAGAGGGGTGCCAGACACCCCTGTATTG
TGCTCTACAAGAGGGCATCGAGCCCTCAGTGGAGATATTTGCCACTGCCATGTGGAAG
AGGTGCCTCCAGCTGCCAGACGACCGGGCAGCCCACGGCTATGGGAGGCCAGCAAGAGG
CTGGCAGGGCTGGGCTGGGAGGATGCTGAACCCGATGAAGACCCCCAGTCTGAGGACTC
AGAGGCCCATCTCTCTAACGACCCCCCACCCTGAGGAGCCCACAGTTCTAACCTTACC
CCAGCCCTCAGAGCTCACCAGATTGTCTAACGATGACGCACCGAATTAGGCTAAAGTTGAG
CCTGAGATCCAGCTCTCT**AA**CCCTCAGGCCAGGATGCTGCCATGGCACTTCATGGCCTT
GAAAACCTCGGATGTGTGAGGCCATGCCCTGGACACTGACGGTTGTGATTTGACCTC
CGTGGTTACTTCTGGGCCCCAAGCTGTGCCCTGGACATCTCTTCTGGTTGAAGGAAT
AATGGGTGATTATTCTTGAGAGTGACAGTAACCCAGATGGAGAGATAGGGTATGCT
AGACACTGTGCTCTCGGAAATTGGATGTAGTATTTCAGGCCACCCCTATTGATTCTG
ATCAGCTCTGGAGCAGAGGCAGGGAGTTGCAATGTGATGCACTGCCAACATTGAGAATTAG
TGAACCTGATCCCTTGCAACCGTCTAGCTAGGTAGTTAACCCATGTTAATGAAGCG
GAATTAGGCTCCCGAGCTAAGGGACTCGCCTAGGGCTCACAGTGAGTAGGAGGAGGGCCTG
GGATCTGAACCCAAGGGCTGAGGCCAGGGCAGTGCCTAAGATGGGTGCTGAGAAGTGA
GTCAGGGCAGGGCAGCTGGTATCGAGGTGCCCATGGAGTAAGGGACGCCCTCCGGCG
ATGCAGGGCTGGGTCACTGTATCTGAAGCCCTCGGAATAAAGCGCGTTGACCGCCAAA
AAAAAAAAAAAAAAA

FIGURE 75

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48227
<subunit 1 of 1, 377 aa, 1 stop
<MW: 40849, pI: 7.98, NX(S/T): 0
MEALLLGAGLLLGYLVLYYNLVKAPPCCGGMGNLRGRTAVVTGANSIGKMTALELARRGAR
VVLACRSQERGEAAAFDLRQESGNNEVIFMALDLASLASVRAFATAFLSSEPRLDILIHNAG
ISSCGRTREAFNLLRVNHIGPFLLTHLLLPCCLKACAPSRRVVVASAHHCRGRRLDFKRLDRP
VVGWRQELRAYADTKLANVLFARRELANQLEATGVTCYAAHPGPVNSELFLRHVPGWLRPLL
PLAWLVLRAPRGGAQTPLYCALQEGIEPLSGRYFANCHVEEVPPAARDDRAAHLWEASKRL
AGLGPGEDAEPDEDPQSEDSEAPSSLSTPHPEEPTVSQPYPSHQSSPDLSKMTHR1QAKVEP
EIQLS
```

Important features:

Signal peptide:

amino acids 1-16

Glycosaminoglycan attachment site.

amino acids 46-49

Short-chain alcohol dehydrogenase family

amino acids 37-49 and 114-124

FIGURE 76

FIGURE 77

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA41404
<subunit 1 of 1, 985 aa, 1 stop
<MW: 105336, pI: 6.55, NX(S/T): 7
MGGMAGDSDPPQILVHPQDQLFQGPGPARMSCQASGQPPPTIRWLLNGQPLSMVPPDPHLLP
DGTLLLLQPPARGHAHDGQALSTDLGVYTCEASNRLGTAWSRGARLSAVLREFQIQPRDM
VAVVGEQFTLECGPPWGHPEPTVWWKDGPLALQPGRHTVSGSLLMARAEKSDGETYMCV
ATNSAGHRESRAARVSIQECPQDYTEPVELLAVRIQLENVTLLNPDAEGPKPRPAVWSWKV
SGPAAPAQSYTALFRTQTAPGGQGAPWAEELLAGWQSAELGGLHWGQDYEFKVRPSSGRARG
PDSNVLLRLPEKVPAPPQEVTLKPGNGTVFVSWVPPPAENHNGIIRGYQVWSLGNDSLPP
ANWTVVGEQTQLEIATHMPGSYCVQVAVTGAGAGEPSRPVCLLLEQAMERATQEPSEHGPW
TLEQLRATLKRPEVIATCGVALWLLLGTAVCIHRRRRARVHLGPLYRYTSEDAILKHRMD
HSDSQWLADTWRSTSGSRDLSSSSSLSSRLGADARDPLDCRRSLLSWDSRSPGVPLLPDTST
FYGSLIAELPSSTPARPSPQVPAVRRLPPQLAQLSSPCSSSDSLSRRLSSPRLSLAPAEA
WKAKKKQELQHANSSPLLRGSHSLELRACELGNRGSKNLSQSPGAVPQALVAWRALGPKLLS
SSNELVTRHLPPAPLFPHETPPTQSQQTQPPVAPQAPSSILLPAAPIPILSPCSPPSPQASS
LSGPSPASSRLSSSSLSSLGEDQDSVLTPPEEVALCLELSEGEETPRNSVSPMPRAPSPTTY
GYISVPTASEFTDMGRTGGVGPKGGVLLCPRPCLTPSEGLANGWGSASEDNAASARA
SLVSSSDGSFLADAHFARALAVAVDSFGFGLEPREADCVFIDASSPPSPRDEIFLTPNLSLP
LWEWRPDWLEDMEVSHTQRLRGMPPWPPDSQISSQRSQLHCRMPKAGASPVDYS
```

Important features:

Transmembrane domain:

amino acids 448-467

N-glycosylation sites:

amino acids 224-227, 338-341, 367-370, 374-377, 658-661 and 926-929

N-myristoylation sites.

amino acids 47-52, 80-85, 88-93, 99-104, 105-110, 181-186, 272-277, 290-295, 355-360, 403-408, 462-467, 561-566, 652-657, 849-854 and 876-881

Phosphotyrosine interaction domain proteins

amino acids 740-753

FIGURE 78

CTCCCACGGTGTCCAGCGCCCAGA**ATG**CGGCTTCTGGTCTGCTATGGGTTGCCTGCTGCT
CCCAGGTTATGAAGCCCTGGAGGGCCCAGAGGAAATCAGCGGTTCGAAGGGACACTGTGT
CCCTGCAGTGCACCTACAGGAAGAGCTGAGGGACCACCGGAAGTACTGGTGCAGGAAGGGT
GGGATCCTCTTCCTCGCTCTGGCACCACATCTATGCAGAAGAAGAAGGCCAGGAGACAAT
GAAGGGCAGGGTGTCCATCCGTGACAGCCGCCAGGAGCTCGCTATTGTGACCCCTGTGGA
ACCTCACCCCTGCAAGACGCTGGGGAGTACTGGTGTGGGTCGAAAAACGGGGCCCCGATGAG
TCTTACTGATCTCTCGTCTTCCAGGACCCCTGTCCTCCCTCCCTCTCCAC
CTTCAGCCTCTGGCTACAAACACGCCTGCAGCCAAAGGAAAGCTCAGCAAACCCAGCCCC
CAGGATTGACTTCTCCTGGGCTCTACCCGGCAGCCACCACAGCCAAGCAGGGGAAGACAGGG
GCTGAGGCCCTCATTGCCAGGGACTTCCAGTACGGCACGAAAGGACTTCTCAGTACAC
AGGAACCTCTCCTCACCCAGCGACCTCTCCTCCTGCAGGGAGCTCCGCCCCCATGCAGC
TGGACTCCACCTCAGCAGAGGACACCAGTCCAGCTCAGCAGTGGCAGCTCTAAGCCAGG
GTGTCCATCCCGATGGTCCGCATACTGGCCCCAGTCCTGGTGTGCTGAGCCTCTGTGAGC
CGCAGGCCCTGATGCCCTCTGCAGCCACCTGTCCTGTGGAGAAAGGAAGCTAACAGGCCA
CGGAGACACAGAGGAACGAGAAAGTTCTGGCTCTCACGCTTGACTGCGGAGGAAAAGGAGCC
CCTTCCCAGGCCCTGAGGGGACGTGATTCGATGCCCTCCACACATCTGAGGAGGA
GCTGGCTCTCGAAGTTGTCTCAGCG**TAGGG**CAGGAGGCCCTGGCCAGGCCAGCAGT
GAAGCAGTATGGCTGGCTGGATCAGCACCGATTCCCGAAAGCTTCCACCTCAGCCTCAGAG
TCCAGCTGCCCGGACTCCAGGGCTCTCCCCACCCCTCCCAAGGCTCTCCTTGCATGTTCA
GCCTGACCTAGAACGCTTGTCAAGCCCTGGAGCCCAGAGCGGTGGCCTGCTCTCCGGCTG
GAGACTGGGACATCCCTGATAGGTTCACATCCCTGGCAGAGTACCAAGGCTGCTGACCCCTA
GCAGGGCCAGACAAGGCTCAGTGGATCTGGCTGAGTTCAATCTGCCAGGAACCTGGC
CTCATGCCAGTGTGGACCCCTGCCCTCCCACTCCAGACCCCCACCTGTCTTCCCCTCCC
TGGCGTCCCTCAGACTTAGTCCCACGGTCTCTGCATCAGCTGGTATGAAGAGGAGCATGCT
GGGGTGAGACTGGGATTCTGGCTCTTGAACCACCTGCATCCAGGCCCTCAGGAAGCCT
GTGAAAACGTGATTCCCTGGCCCCACCAAGACCCACCAAAACCATCTCTGGGCTTGGTGCAG
GACTCTGAATTCTAACATGCCAGTGACTGTCGCACTTGAGTTGAGGGCCAGTGGCCTG
ATGAACGCTCACACCCCTCAGCTTAGAGTCTGCATTGGCTGTGACGTCTCACCTGCC
CAATAGATCTGCTCTGCGACACCAGATCCACGTGGGACTCCCTGAGGCCTGCTAAG
TCCAGGCCCTGGTCAGGTCAAGTGCACATTGCAGGATAAGCCCAGGACCGGCACAGAAGTGG
TTGCCTTNCATTGCCCTCCCTGGNCATGCCTTCTGCCTTGGAAAAAATGATGAAGA
AAACCTTGGCTCCTCCTGGAAAGGGTTACTTGCCTATGGGTCTGGTGGCTAGAGA
GAAAAGTAGAAAACCAGAGTGCACGTAGGTGCTAACACAGAGGAGAGTAGGAACAGGGCGG
ATACCTGAAGGTGACTCCGAGTCCAGGCCCTGGAGAAGGGTGGGGTGGTAAAGTA
GCACAACACTATTTCCTTCCATTATTATTGTTTAAGACAGAACTCGTGCT
GCTGCCAGGCTGGAGTGCAGTGGCACGATCTGCAAACCTCCGCTCTGGTTCAAGTGATT
CTTCTGCCTCAGCCTCCGAGTAGCTGGATTACAGGCACGCACCACACCTGGCTAATT
TTTGTACTTTAGTAGAGATGGGTTTCACCATGTTGCCAGGCTGGCTTGAACCTTGAC
CTCAAATGAGCCTCTGCTCAGTCTCCAAATTGCCGGATTACAGGCATGAGCCACTGTG
TCTGCCCTATTCCCTTAAAAGTGAATTAAGAGTTGTTAGTATGCAAACACTGGAAAG
ATGGAGGAGAAAAGAAAAGGAAGAAAAAATGTCAACCATAGTCTCACCAGAGACTATCAT
TATTCGTTTGTGACTTCCTCACTTTCTTCTTCACATAATTGCCGGTGTCTT
TTTACAGAGCAATTATCTGTATATACAACCTTGTATCCTGCCCTTCCACCTTATCGTTCC
ATCACTTATTCCAGCACTCTGTGTTACAGACCTTTATAAAATAAAATGTTCATCA
GCTGCATAAAAAAAAAAAAAAA

FIGURE 79

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44196
<subunit 1 of 1, 332 aa, 1 stop
<MW: 36143, pI: 5.89, NX(S/T): 1
MRLLVLLWGCLLPGYEALEGPEEISGFEGDTVSLQCTYREELRDHRKYWCRKGGILFSRCS
GTIYAAEEEGQETMKGRVSIRDSRQELSLIVTLWNLTQDAGEYWCGVEKRGPDSSLISLFV
FPGPCCPPSPSPTFQPLATTRLQPKAKAQQTQPPGLTSPGLYPAATTAKQGKTGAEAPPLPG
TSQYGHERTSQYTGTSPHPATSPPAGSSRPPMQLDSTSADTSPALSSGSSKPRVSIPMVRI
LAPVLVLLSLLSAAGLIAFCSHLLLWRKEAQQATETQRNEKFWLSRLTAEEKEAPSQAPEGD
VISMPPLHTSEEELGFSKFVSA
```

Important features:

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 248-269

N-glycosylation site.

amino acids 96-99

Fibrinogen beta and gamma chains C-terminal domain.

amino acids 104-113

Ig like V-type domain:

amino acids 13-128

FIGURE 80

TTGTGACTAAAGCTGGCTAGCAGGCCAGGGAGTGCAGCTGCAGGCAGGGGTGGCAGGA
GCCGCAGAGCCAGAGCAGACAGCCGAGAAACAGGTGGACAGTGTGAAAGAACCAAGTGGTCTC
GCTCTGTTGCCAGGCTAGAGTGTACTGGCGTGTACAGCTCACTGCAGCCTCAGACTCCT
GGACTTGAGAAATCCTCCTGCCTTAGCCTCCTGCATATCTGGACTCCAGGGTGCACTC
GCCCTGTTCTCCTCTGTGAGTGGACACGGAGGCTGGTGGACTGCCTGTCATCCCAA
AGCTCAGCTCTGAGCCAGAGTGGTGGCTCCACCTCTGCCGCCATAGAACCCAGGAG
CAGGGCTCTCAGAAGGCGGTGGTGCCTCAGCTGGATCATGTTGTTGGCCCTGGTCTGCTGC
TCAGCTGCCCTGCTACCCCTCAGTGAGGCCAAGCTCTACGGTCGTTGTGAACTGCCAGAGTG
CTACATGACTTCGGGCTGGACGGATACCGGGGATACAGCCTGGCTGACTGGGCTGCCCTGC
TTATTCACAAGCGGTTCAACGCAGCTGCTTGACTACGAGGCTGATGGGAGCACCAACA
ACGGGATCTTCCAGATCAACAGCCGGAGGTGGTCAGCAACCTCACCCGAACGTCCCCAAC
GTGTGCCGGATGTACTGCTCAGATTGTTGAATCCTAATCTAAGGATACCCTATCTGTGC
CATGAAGATAACCCAAGAGCCTCAGGGTCTGGTTACTGGGAGGCCTGGAGGCATCACTGCC
AGGGAAAAGACCTCACTGAATGGGTGGATGGCTGTGACTTCTAGGATGGACGGAACCATGCA
CAGCAGGCTGGAAATGTGGTTGGTCTGACCTAGGCTGGGAAGACAAGCCAGCGAATA
AAGGATGGTTGAACGTGAAA

FIGURE 81

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52187
<subunit 1 of 1, 146 aa, 1 stop
<MW: 16430, pI: 5.05, NX(S/T): 1
MLLALVCLLSCLLPSEAKLYGRCELARVLHDFGLDGYRGYSLADWVCLAYFTSGFNAALD
YEADGSTNNGIFQINSRRWCSNLTPNVPNVCRMYCSDLLNPNLKDTVICAMKITQEPOGLGY
WEAWRHHCQGKDLTEWVDGCDF
```

Important features:

Signal peptide:

amino acids 1-18

N-myristoylation site.

amino acids 67-72

Homologous region to Alpha-lactalbumin / lysozyme C proteins.

amino acids 34-58 (catalytic domain), 111-132 and 66-107

FIGURE 82

AGCCGCTGCCCGGGCGCCCGCGGGCACCATGAGTCCCCGCTCGTGCCTGCGTTG
GCTGCGCCTCCTCGTCTCGCCGTCTCTCAGCCGCCGAGCAACTGGCTGTACCTGGCA
AGCTGTCGTCGGTGGGGAGCATCTCAGAGGAGGAGACGTGCGAGAAACTCAAGGGCTGATC
CAGAGGCAGGTGCAGATGTGCAAGCGGAACCTGGAAGTCATGGACTCGGTGCCCGGGTGC
CCAGCTGCCATTGAGGAGTGCCAGTACAGTTCCGGAACCGGCCTGGAACGTGCTCCACAC
TCGACTCCTGCCGTCTCGGCAAGGTGGTACGCAAGGGACTCGGGAGGCGGCCCTCGT
TACGCCATCTCTCGGCAGGTGTGGCCTTGCACTGACGCGGGCGTGCAGCAGTGGGAGCT
GGAGAAGTGCAGGCTGTGACAGGACAGTCAGGGTCAGCCCACAGGGCTTCCAGTGGTCAG
GATGCTCTGACAACATCGCCTACGGTGTGGCCTCTCACAGTCGTTGTGGATGTGCAGGAG
AGAAGCAAGGGGCCTCGTCCAGCAGGCCCTCATGAACCTCCACAACAATGAGGCCAG
GAAGGCCATCCTGACACACATGCCGGTGGAAATGCAAGTGCACGGGTGTCAGGCTCCTGTG
AGGTAAAGACGTGCTGGCGAGCCGTGCCGCCCTCCGCCAGGTGGTCACGCACGTGAAGGAG
AAGTTGATGGTGCCACTGAGGTGGAGCCACGCCGCGTGGCTCCTCCAGGGCACTGGTAC
ACGCAACGCACAGTTCAAGCCGCACACAGATGAGGACCTGGTACTTGGAGCCTAGCCCCG
ACTTCTGTGAGCAGGACATGCGCAGCGCGTGCTGGCACGAGGGCCGCACATGCAACAAG
ACGTCCAAGGCCATCGACGGCTGTGAGCTGCTGTGCTGGCCGCCCTCCACACGGCGCA
GGTGGAGCTGGCTGAACGCTGCAGCTGCAAATTCCACTGGTGTGCTCGTCAAGTGCAGG
AGTGCCAGCGGCTCGTGGAGTTGCACACGTGCCGATGACGCCCTGCCTAGCCCTGCGCCGGC
AACCCACCTAGTGGCCCAAGGAAGGCCATAATTAAACAGTCTCCCACCACTACCCCAAGA
GATACTGGTTGTATTTTGTCTGGTTGGGTTGGTCCTCATGTTATTATTGCCGAA
ACCAGGCAGGCAACCCCAAGGGCACCAACCAGGGCTCCCCAAAGCCTGGCCTTGTGGCT
GCCACTGACCAAAAGGGACCTTGCTCGTGCCTGGCTGCCGATGTGGCTGCCACTGACCA
CTCAGTTGTTATCTGTGTCCGTTTCTACTGCAAGACCTAAGGTGGAGTAACAAGGAGTAT
TACCACCACATGGCTACTGACCGTGTCACTGGGAAGAGGGGGCTTATGGCAGGGAAAATA
GGTACCGACTTGATGGAAGTCACACCCCTCTGGAAAAAAAGAAACTCTTAACCTCCAGCACACA
TACACATGGACTCCTGGCAGCTTGAGCCTAGAACCCATGTCTCTCAAATGCCCTGAGAAAGG
GAACAAAGCAGATACCAGGTCAAGGGCACCAAGGTTCACTTCCAGCCCTACATGGACAGCTAGA
GGTCGATATCTGTGGTCCTCCAGGCAAGAACAGGGAGATGAGAGCAAGAGACGACTGAA
GTCCCACCCCTAGAACCCAGCCTGCCCAAGCCTGCCCTGGGAAGAGGAAACTTAACCACTCC
CCAGACCCACCTAGGCAGGCATATAGGCTGCCATCCTGGACCAGGGATCCCGGCTGTGCCCT
TGCAGTCATGCCGAGTCACCTTCACAGCGCTGTTCTCCATGAAACTGAAAAACACACAC
AC
GAGAGGGAGGAAAGGGCTGTGCCTTGCACTGCCCCAGTCACCTTCACAGCACTGTCCTC

FIGURE 83

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48328
<subunit 1 of 1, 351 aa, 1 stop
<MW: 39052, pI: 8.97, NX(S/T): 2
MSPRSCLRSLLVFAVFSAAASNWLYLAKLSSVGSISEEETCEKLKGLIQRQVQMCKRNLE
VMDSVRRGAQLAIEECQYQFRNRRWNCSTLDSLPVFGKVVTQGTREAAFVYAISSAGVAFAV
TRACSSGELEKCGCDRTVHGVSPQGFQWSGCSDNIAYGVAFSQSFVDVRERSKGASSRALM
NLHNNEAGRKAILTHMRVECKCHGVSGSCEVKTCWRADVPPFRQVGHALKEKFDGATEVEPRR
VGSSRALVPRNAQFKPHTDEDLVYLEPSPDFCEQDMRSGVLGTRGRTCNKTSKAIDGCELLC
CGRGFHTAQVELAERCSCKFHWCCFVKCRQCQRLVELHTCR
```

Important features:

Signal peptide:

amino acids 1-22

N-glycosylation sites.

amino acids 88-91 and 297-300

Wnt-1 family signature.

amino acids 206-215

Homologous region to Wnt-1 family proteins

amino acids 183-235, 305-350, 97-138, 53-92 and 150 -174

FIGURE 84

CGGACGCGTGGCGGACGCGTGGCGGACGCGTGGCGGACGCGTGGGCTGGGTGCCTGCAT
CGCCATGGACACCAGGTACAGCAAGTGGGGCGGAGCTCGAGGAGGTCCCCGGAGGGC
CCTGGGGACGCTGGGTGCACTGGAGCAGGAGACCCCTTTCTGGCCCTGGCTGTCCGGTC
ACACAGTCCTTGGCTGTGATTCTGAGTATCCTATTGTCCAAGGCCTCCACGGAGCGCGC
GGCGCTGCTTGACGGCCACGACCTGCTGAGGACAAACGCCTCGAAGCAGACGGCGCGCTGG
GTGCCCTGAAGGAGGAGGTGGAGACTGCCACAGCTGCTGCTGGGGACGCAGGCGCAGCTG
CAGACCACGCGCGGGAGCTTGGGGAGGCGCAGGCGAAGCTGATGGAGCAGGAGAGGCCCT
GCGGGAACTGCGTGAGCGCGTGACCCAGGGCTTGGCTGAAGCCGGCAGGGCCGTGAGGACG
TCCGCACTGAGCTGTTCCGGCGCTGGAGGCCGTGAGGCTCCAGAACAACTCCTGCGAGCCG
TGCCCCACGTCGTGGCTGTCCCTGAGGGCTCCTGCTACTTTCTGTGCCAAAGACGAC
GTGGCGGGCGCAGGATCACTGCGCAGATGCCAGCGCACCTGGTGATCGTTGGGGCC
TGGATGAGCAGGGCTTCCTCACTCGAACACCGCGTGGCGTGGTTACTGGCTGGGCCTGAGG
GCTGTGCCATCTGGCAAGGTTAGGGCTACCAAGCTGGGTGGACGGAGTCTCTCAGCTT
CAGCCACTGGAACCAGGGAGAGCCCAATGACGCTTGGGGCGCGAGAACTGTGTCATGATGC
TGCACACGGGCTGTGGAACGACGCACCGTGTGACAGCGAGAAGGACGGCTGGATCTGTGAG
AAAAGGCACAACTGCTGACCCGCCAGTGCCCTGGAGGCCGCCATTGCAGCATGTCGA
TCCTGGGGCTGCTCACCTCCCTGGCTCCTGGAGCTGATTGCCAAAGAGTTTTCTTCCT
CATCCACCGCTGCTGAGTCTCAGAAACACTTGGCCAACATAGCCCTGTCCAGCCAGTGC
TGGGCTCTGGACCTCCATGCCGACCTCATCTAACTCCACTCACGCAGACCCAACTAAC
TCCACTAGCTCCAAAATCCCTGCTCCTGCGTCCCGTGATATGCCTCCACTCTCCCTAA
CCAAGGTTAGGTGACTGAGGACTGGAGCTGGAGCTGGTTGGTTCTCGCATTTCACCAAACTGGA
AGCTTTTGCAAGCTGAGGAAGCATCAATAAATATTGAGAAATGAAAAAA

FIGURE 85

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56352
<subunit 1 of 1, 293 aa, 1 stop
<MW: 32562, pI: 6.53, NX(S/T): 2
MDTTRYSKWGGSEEVPGGPWGRVWHSRRPLFLALAVLVTTLWAVILSILLSKASTERAA
LLDGHDLLRTNASKQTAALGALKEEVGDCHSCCSGTQAQLQTTRAELGEAQAKLMEQESALR
ELRERVTQGLAEAGRGRGREDVRTELFRALLEAVRLQNNSCEPCPSTSMLSFEGLSCYFFSVPKTTW
AAAQDHCADASAHLVIVGGLDEQGFLTRNTRGRGYWLGLRAVRHLGKVQGYQWVDGVSLSF
HWNQGEPNDAWGRENCVMMMLHTGLWNDAPCDSEKDGWICEKRHNC
```

Important features:

Type II transmembrane domain:

amino acids 31-54

N-glycosylation sites.

amino acids 73-76 and 159-162

Leucine zipper pattern.

amino acids 102-123

N-myristoylation sites.

amino acids 18-23, 133-138 and 242-247

C-type lectin domain signature.

amino acids 264-287

FIGURE 86

GCCAGGGGAAGAGGGTATCCGACCCGGGAAGGTCGCTGGCAGGGAGTTGGAAAGCG
GCAGCCCCCGCCGCCCGCAGCCCTTCTCCTCCTTCTCCACGTCTATCTGCCTCTCG
CTGGAGGCCAGGCCGTGCAGCATCGAAGACAGGAGGAACCTGGAGCCTATTGCCGGCCGG
GGCGCCGGCCTCGGGCTTAAATAGGAGCTCCGGCTCTGGCTGGACCCGACCGCTGCCGGC
CGCGCTCCGCTGCTCCTGCCGGTGAGGAAAACCCCAGCCGGCCGCCCTGGCAAG
GCCCTCTGCGCTCTCCTGCCACTCTCGCGCCGGCAGCCTCTGGGGAGAGTC
CATCTGTTCCGCCAGAGCCCCGGCAAATACAGCATCACCTCACGGCAAGTGGAGCCAGA
CGGCCTTCCCCAAGCAGTACCCCTGTTCCGCCCCCTGCGCAGTGGCTTCGCTGGGG
GCCGCGCATAGCTCCGACTACAGCATGTGGAGGAAGAACAGTACGTCAGTAACGGCTGCG
CGACTTGGAGCGCGGGAGGCCCTGGCGCTGATGAAGGAGATCGAGGCGGGGGAGG
CGCTGCAGAGCGTGCACGAGGTGTTTGGCGCCCGTCCCCAGCGGACCCGGCAGACG
TCGGCGGAGCTGGAGGTGCAGCGCAGGCACTCGCTGGTCTCGTTGTGGTGCATCGGCC
CAGCCCCGACTGGTCGTGGCGTGGACAGCCTGGACCTGTGCGACGGGACCGTTGGCGGG
AACAGGCAGCGCTGGACCTGTACCCCTACGACGCCGGACGGACAGCGGCTTCACCTCTCC
TCCCCCAACTCGCCACCATCCCGAGGACACGGTACCGAGATAACGTCCTCCTCTCCAG
CCACCCGGCCAACTCCTCTACTACCCCGGGCTGAAGGCCCTGCCCTCCATGCCAGGGTGA
CACTGCTGGCTGCGACAGAGCCCCAGGGCTTCATCCCTCCGCCAGTCCTGCCAGC
AGGGACAATGAGATTGTAGACAGCGCTCAGTTCCAGAAACGCCGGACTGCGAGGTCTC
CCTGTGGTCGTCTGGGACTGTGGAGGCCACTGTGGAGGCTGGACCAAGAGCAGGA
CTCGCTACGTCCGGTCCAGCCGCAACACGGGAGCCCTGCCCGAGCTCGAAGAAGAG
GCTGAGTGCCTGATAACTGCGTCTAAGGACCAGAGCCCCGAGCCCTGGGGCCCCCG
GAGCCATGGGTGTCGGGGCTCTGTGCAGGCTATGCTGCAGGCTCATGCTGCAGGCGGCCAGGG
GGTTTCGCGCTGCTCCTGACCGCGGTGAGGCCGCGCCACCATCTGCACTGAAGGGCCCT
CTGGTGGCCGGCACGGCATTGGAAACAGCCTCCTCCTTCCAACCTTGCTTCTAGGG
CCCCCGTGTCCGTCTGCTCTCAGCCTCCTCCTGCAGGATAAAAGTCATCCCCAAGGCTC
CAGCTACTCTAAATTATGTCCTTATAAGTTATTGCTGCTCCAGGAGATTGTCTTCATCG
TCCAGGGGCCTGGCTCCACGTGGTTGCAGATACTCAGACCTGGTGCCTAGGCTGTGCTG
AGCCCACCTCCGAGGGCGCATCCAAGCGGGGCCACTTGAGAAGTGAATAAATGGGCAG
TTTCGGAAGCGTCAGTGTTCATGTTATGGATCTCTGCGTTGAATAAAGACTATCTCT
GTTGCTCACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 87

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53971
><subunit 1 of 1, 331 aa, 1 stop
><MW: 35844, pI: 5.45, NX(S/T): 2
MENPSPAAALGKALCALLLATLGAAGQPLGGESICSARAPAKYSITFTGKWSQTAFPKQYPL
FRPPAQWSSLGAAHSSDYSMWRKNQYVSNGLRDFAERGEAWALMKEIEAAGEALQSVHEVF
SAPAVPSGTGQTSAELEVQRRHSLVSFVVRIVPSPDWFGVDSLDLCDGDRWREQAALDLYP
YDAGTDSGFTFSSPNFATIPQDTVTEITSSSPSHPANSFYYPRLKALPPIARVTLLRLRQSP
RAFIPPPAPVLPSPRDNEIVDSASVPETPLDCEVSLWSSWGLCGGHCGRGLGTSRTRYVRVQPA
NNGSPCPELEEEAECVPDNCV
```

Important features:

Signal peptide:

amino acids 1-26

FIGURE 88

GGCGGGCGTCCGTGAGGGGCTCCTTGGCAGGGTAGTGTGTTGGTGTCCCTGTCTGCGTGA
TATTGACAAACTGAAGCTTCCTGCACCACTGGACTTAAGGAAGAGTGTACTCGTAGGCGGA
CAGCTTAGTGGCCGGCCGGCGCTCTCATCCCCGTAAGGAGCAGAGTCCTTGTACTGAC
CAAGATGAGCAACATCTACATCCAGGAGCCTCCCACGAATGGGAAGGTTTATTGAAAACTA
CAGCTGGAGATATTGACATAGAGTTGTGGTCAAAGAAGCTCCTAAAGCTTGCAGAAATT
ATCCAACCTTGTTGGAAGCTTATTATGACAATACCATTTCATAGAGTTGTGCCTGGTT
CATAGTCCAAGGCGGAGATCCTACTGGCACAGGGAGTGGTGGAGAGTCTATCTATGGAGCGC
CATTCAAAGATGAATTCATTACGGTTGCCTTAATCGGAGAGGACTGGTTGCCATGGCA
AATGCTGGTTCTCATGATAATGGCAGCCAGTTTCTTCACACTGGGTGAGCAGATGAACT
TAACAATAAGCATACCATTGGAAAGGTTACAGGGGATACAGTATATAACATGTTGCGAC
TGTCAAGTAGACATTGATGATGACGAAAGACCACATAATCCACACACAAAATAAAAGCTGT
GAGGTTTGTAAATCCTTTGATGACATCATTCCAAGGGAAATTAAAAGGCTGAAAAAAGA
GAAACCAGAGGAGGAAGTAAAGAAATTGAAACCCAAAGGCACAAAAAATTAGTTACTTT
CATTGGAGAGGAAGCTGAGGAAGAAGAGGGAGTAAATCGAGTTAGTCAGAGCATGAAG
GGCAAAAGCAAAAGTAGTCATGACTTGCTTAAGGATGATCCACATCTCAGTTCTGTTCCAGT
TGTAGAAAGTGAAGAAAGGTGATGCACCAAGATTAGTTGATGATGGAGAAGATGAAAGTGCAG
AGCATGATGAATATATTGATGGTGTGAAAGAACCTGATGAGAGAAAGAATTGCCAAAAAA
TTAAAAAAGGACACAAGTGCATGAACTGTTAAATCAGCTGGAGAAGGAGAAGTGGAGAAGAAATC
AGTCAGCCGCAGTGAAGAGCTCAGAAAAGAAGCAAGACAATTAAAACGGGAACCTCTAGCAG
CAAAACAAAAAAAGTAGAAATGCAGCAAAACAAGCAGAAAAAGAAGTGAAGAGGAAGAA
GCCCTCCAGATGGTGTGCTTGCAGAACACAGAGTGAAGCTTGAG
GAAGCAACAGTCAGGAAACTTCCCAGGAAGATCAGACCCCTGCACTGCTGAACCAAGT
TTAAATCTAAACTCACTCAAGCAATTGCTGAAACACCTGAAAATGACATTCTGAAACAGAA
GTAGAAGATGATGAAGGATGGATGTCACATGTACTTCAGTTGAGGATAAAAGCAGAAAAGT
GAAAGATGCAAGCATGCAAGACTCAGATAACATTGAAATCTATGATCCTCGGAATCCAGTGA
ATAAAAGAAGGAGGAAAGAAAGCAAAAGCTGATGAGAGAGAAAAAGAAAGAAGATAAAA
GAGAATAATGATAACCAGAACTTGCTGGAAATGTGCCTACAATGGCCTGTAACAGCCATTG
TTCCCAACAGCATCACTTAGGGGTGTGAAAAGAAGTATTTGAACTGTTGTCTGGTTTG
AAAAACAATTATCTGTTGCAAATTGTGGAATGATGTAAGCAAATGCTTTGGTTACTGG
TACATGTGTTTCTAGCTGACCTTTATATTGCTAAATCTGAAATAAAACTTCCCT
TCCACAAAAAA

FIGURE 89

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50919
><subunit 1 of 1, 472 aa, 1 stop
><MW: 53847, pI: 5.75, NX(S/T): 2
MSNIYIQEPPPTNGKVLKTTAGDIDIELWSKEAPKACRNFQLCLEAYYDNTIFHRVVPGFI
VQGGDPTGTGSGGESIYGAPFKDEFHSRLRFNRRGLVAMANAGSHDNGSQFFFTLGRADELN
NKHTIFGKVTGDTVYNMLRLSEVDIDDDERPHNPHKIKSCEVLNFNPFDIIPREIKRLKEK
PEEEVKKLKPKGTKNFSLLSFGEAAEEEEEVNRVSQSMKGKSKSSHDLKDDPHLSSVPVV
ESEKGDAPDLVDDGEDESAEHDEYIDGDEKNLMRERIAKKLKDTSANVKSAGEGEVEKKSV
SRSEELRKEARQLKRELLAAKQKKVENAAKQAEKRSEEEAPPDGAVAAYRREKQKYEALRK
QQSKKGTREDQTLALLNQFKSKLTQAIATEPNDIPETEVEDDEGWMHVLQFEDKSRKVK
DASMQDSDTFFIYDPRNPVNKRREESKKLMREKKERR
```

Important features:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 109-112 and 201-204

Cyclophilin-type peptidyl-prolyl cis-trans isomerase signature.

amino acids 49-66

Homologous region to Cyclophilin-type peptidyl-prolyl cis-trans isomerase

amino acids 96-140, 49-89 and 22-51

FIGURE 90

CGCCGCCGTTGGGCTGGAAGTTCCCGCCAGGTCCGTGCCGGCGAGAGAGATGCTGCCCGG
CCCGCCTCGGCTTGAGGGAGAGAAGTGTCCCAGACCCATTGCGCTGACGGCGTCG
AGCCCTGGCCAGAC**A**TGTCCACAGGGTTCTCCTCGGGACTCTGGGCTCCACCA
GTGGCCGCCGGCGGGACCAGCACAGGCGGCGTTCTCCTCGGAACGGGAACGTCTAGCAA
CCCTTCTGTGGGGCTCAATTGGAAATCTTGAAGTACTTCAACTCCAGCAACTACATCTG
CTCCTCAAGTGGTTTGAACCGGGCTCTTGGATCTAACCTGCCACTGGGTTCACTCTA
GGAGGAACAAATACAGGTGCCTGCACACCAAGAGGCCTCAAGTGGTCAACAAATATGGAAC
CCTGCAAGGAAAACAGATGCATGTGGGAAGACACCCATCCAAGTCTTTAGGAGTCCCCT
TCTCCAGACCTCCTCTAGGTATCCTCAGGTTGCACCTCCAGAACCCCCGGAGCCCTGGAAA
GGAATCAGAGATGCTACACCTACCCGCCTGGATGGAGTCTCGCTCTGCGCCAGGCTGGAG
TGCAGTGGCACGATCTGGCTCACTGCAACCTCCGCCTCCGGTTCAAGCGAGTCTCCTGC
CTCAGCCTCTGAGTGTCTGGGCTACAGGTGCCTGCAGGAGTCTGGGGCCAGCTGGCCTCG
ATGTACGTCAGCACGGAAACGGTACAAGTGGCTGCGCTTCAGCGAGGACTGTCTGTACCT
GAACGTGTACGCCGCCGGCGCGCCGGGATCCCCAGCTGCCAGTGTATGGTCTGGTCC
CGGGAGGGCGCTTCATCGTGGCGCTGCTTCTCGTACGAGGGCTCTGACTTGGCCGCCCG
GAGAAAGTGGTGTGGTTCTGCAGCACAGGCTGGCATCTCGGCTTCTGAGCACGGA
CGACAGCCACGCCGGAACTGGGGCTGGACCAGATGGCGGCTCTGCGCTGGTGC
AGGAGAACATCGCAGCCTCGGGGAGACCAGGAATGTGACCCCTGTTGGCCAGTCGGCG
GGGCCATGAGCATCTCAGGACTGATGATGTCACCCCTAGCCTGGGTCTTCCATGGGC
CATTTCCAGAGTGGCACCGCGTTATTAGACTTTCATCACTAGTAACCCACTGAAAGTGG
CCAAGAAGGTTGCCACCTGGCTGGATGCAACCACACAGCACAGATCCTGGTAAACTGC
CTGAGGGCACTATCAGGGACCAAGGTGATGCGTGTCCAACAAGATGAGATTCTCCA
GAACTTCCAGAGAGACCCGGAAGAGATTATCTGGTCCATGAGCCCTGGTGGATGGTGTGG
TGATCCCAGATGACCCCTTGGCTCCTGACCCAGGGGAAGGTTCATCTGTGCCCTACCTT
CTAGGTGTCAACAACCTGGAATTCAATTGGCTTTCGCTTATAATATCACCAAGGAGCAGGT
ACCACTTGTGGAGGAGTACCTGGACAATGTCAATGAGCATGACTGGAAGATGCTACGAA
ACCGTATGATGGACATAGTCAAGATGCCACTTCGTGTATGCCACACTGCAGACTGCTCAC
TACCAACGAGAAACCCAAATGATGGAATCTGCCCTGCTGGCCACGCTACAACAAGGATGAA
AAGTACCTGCAGCTGGATTTCACACAAGAGTGGC**A**TGAAGCTCAAGGAGAAGAAGATGGC
TTTTGGATGAGTGTACCAAGACTGAGAAGCAGAGGAATTCTAAGGGTGGC
TATGCAGGAAGGAGGCCAAAGAGGGTTGCCACCATCCAGGCCCTGGGAGACTAGCCA
TGGACATACTGGGACAAGAGTTCTACCCACCCAGTTAGAACTGCAGGAGCTCCCTGCT
GCCTCCAGGCCAAAGCTAGAGCTTTGCCTGTTGTGGACCTGCACTGCCCTTCCAGCC
TGACATCCCATGATGCCCTCACTTCAGTGTGACATCCAGTTAGGCCAGGCCCTGTCAAC
ACCACACTGTGCTCAGCTCCAGCCTCAGGACAACCTTTTCCCTTCAATCCT
CCCACCTCAATGTCTCCTGTGACTCCTCTTATGGGAGGTGACCCAGACTGCCACTGC
CCCTGTCACTGCACCCAGCTGGCATTACCATCCATCCTGCTCAACCTGTTCTGTCTGT
TCACATTGGCCTGGAGGCCTAGGGCAGGTGTGACATGGAGCAAACCTTTGGTAGTTGGGA
TCTTCTCTCCCACCCACACTTATCTCCCCAGGGCACTCCAAAGTCTATACACAGGGTGG
TCTCTCAATAAAGAAGTGGATTAGAAAAAA

FIGURE 91

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44179
<subunit 1 of 1, 545 aa, 1 stop
<MW: 58934, pI: 9.45, NX(S/T): 4
MSTGFSFGSGTLGTTVAAGGTSTGGVFSFGTGTSSNPSVGLNFGNLGSTSTPATTsapSSG
FGTGLFGSKPATGFTLGGTNTGALHTKRPQVVTKYGTLQGKQMHVGKTPIQVFLGVFSPRPP
LGILRFAPPEPPEPKGIRDAATTYPPGWSLALSPGWSAVARSRLTATSASRVQASLLPQPLS
VWGYRCLQESWGQLASMYVSTRERYKWLRFSEDCLYLNVYAPARAPGDPQLPVMVWFPGGAF
IVGAASSYEGSDLAAREKVVLVFLQHRLGIFGFLSTDDSHARGNWGLLDQMAALRWVQENIA
AFGGDPGNVTLFGQSAGAMSIISGLMMSPLASGLFHRAISQSGTALFRLFITSNPLKVAKKVA
HLAGCNHNSTQILVNCLRALSGTKVMRVSNKMRFLQLNFQRDPEEIWSMSPVVVDGVVIPDD
PLVLLTQGKVSSVPYLLGVNNLEFNWLLPYNITKEQVPLVVEYLDNVNEHDWKMLRNRMMD
IVQDATFVYATLQTAHYHRETPMMGICPAGHATTRMKSTCSWILPQEWA
```

Important features:

Signal peptide:

amino acids 1-29

Carboxylesterases type-B serine active site.

amino acids 312-327

Carboxylesterases type-B signature 2.

amino acids 218-228

N-glycosylation sites.

amino acids 318-321, 380-383 and 465-468

FIGURE 92

GAGAACAGGCCTGTCAGGCAGGCCCTGCCCTCCTATGCGGAGATGCTACTGCCACTGCT
GCTGTCCTCGCTGGCGGGTCCCAGGCATGGATGGGAGATTCTGGATACGAGTCAGG
AGTCAGTGATGGTGCAGGGCCTGTGCATCTCTGCCCCCTGCTCTTCTCCTACCCCCGA
CAAAGACTGGACAGGGTCTACCCAGCTTATGGCTACTGGTTCAAAGCAGTGAGACAAAC
CAAGGGTGCCTGTGGCCACAAACCACAGAGTCAGAGGGAAATGAGCAGCCGGGCC
GATTCCAGCTCACTGGGATCCCAGCAAGGGAACTGCTCCTGGTATCAGAGACGCCAG
ATGCAGGATGAGTCACAGTACTTCTTCGGGTGGAGAGAGGAAGCTATGTGACATATAATT
CATGAACGATGGTTCTTCTAAAAGTAACAGTGCTCAGCTCACGCCAGACCCAGGACC
ACAACACCGACCTCACGCCATGTGGACTCTCCAGAAAGGGTGTGAGCGCACAGAGGACC
GTCGACTCCGTGTGGCTATGCCCGAGAGACCTGTTATCAGCATTACGTGACAACAC
GCCAGCCCTGGAGCCCCAGCCCCAGGGAAATGTCCCACCTGGAAGCCAAAAGGCCAGT
TCCTGCGGCTCCTGTGCTGACAGCCAGCCCCCTGCCACACTGAGCTGGGTCTGCAG
AACAGAGTCCTCTCCTCGTCCCACCCCTGGGCCCTAGACCCCTGGGCTGGAGCTGCCGG
GGTGAAGGCTGGGATTCAAGGGCCTACACCTGCCAGCGGAGAACAGGCTTGGCTCCAGC
AGCGAGCCCTGGACCTCTGTGCACTATCCTCCAGAGAACCTGAGAGTGATGGTTCCAA
GCAAACAGGACAGTCCTGGAAAACCTGGAACGGCACGTCTCCAGTACTGGAGGGCA
AAGCCTGTGCCCTGGTCTGTGTCACACACAGCAGCCCCCAGCCAGGCTGAGCTGGACCCAGA
GGGGACAGGTTCTGAGCCCCCTCCAGGCCCTCAGACCCCCGGGTCTGGAGCTGCCCTGGGTT
CAAGTGGAGCACGAAGGAGAGTTCACCTGCCACGCTCGCACCCACTGGCTCCAGCACGT
CTCTCTCAGCCTCTCGTGCACTATAAGAAGGGACTCATCTCACGGCATTCTCCAACGGAG
CGTTCTGGGAATGGCATCACGGCTCTTCTTCCTGCTGGCCCTGATCATCATGAAG
ATTCTACCGAAGAGACGGACTCAGACAGAAACCCGAGGCCAGGTTCTCCGGCACAGCAC
GATCCTGGATTACATCAATGGTCCCGACGGCTGGCCCCCTGGCTCAGAAGGGAAATCAGA
AAGCCACACCAAACAGTCCTGGACCCCTCTCCACCGGTGCTCCCTCCCCAGAATCAAAG
AAGAACCAAGAAAAGCAGTATCAGTTGCCAGTTCCAGAACCCAAATCATCCACTCAAGC
CCAGAACATCCCAGGAGAGCCAAGAGGAGCTCATTATGCCACGCTCAACTCCAGGGTCA
GACCCAGGCCTGAGGCCGGATGCCAACGGCACCCAGGGATTATGAGAAGTCAAGTTC
CAATGAGGGTCTTTAGGCTTAGGACTGGACTTCGGCTAGGGAGGAAGGTAGAGTAAGAG
GTTGAAGATAACAGAGTGAAAGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT
CT
CCAGCACTTGGGAGGTGGCAGATGCCCTGAGGTGGGAGTTCGAGGACCAGCCTG
GCCAAGTGGTGAACACCCGCTCTACTAAAAAATACAAAAATTAGCTGGCATGGTGGCAGG
CGCCTGTAATCCTACCTACTGGGAAGCTGAGGCAGGAGAACATCACTGAACCTGGGAGACGG
AGGTTGCAGTGAGCCAAGATCACACCATTGCACGCCAGCCTGGCAACAAAGCGAGACTCCA
TCTCAAAAAAAATCCTCAAATGGGTTGGGTGTTGTAATCCCAGCACTTGGGAGGCTA
AGGTGGGTTGGATTGCTTGAGGCCAGGAGTTGAGACAGCCTGGCAACATGGTAAACCC
ATCTCTACAAAAAATACAAAACATAGCTGGGCTTGGTGTGCTGCTGAGTCCCAGCTGT
CAGACATTAAACCAGAGCAACTCCATCTGGAAATAGGAGCTGAATAAAATGAGGCTGAGACC
TACTGGGCTGCATTCTCAGACAGTGGAGGCATTCTAAGTCACAGGATGAGACAGGGAGGTCCG
TACAAGATACAGGTCAAAAGACTTTGCTGATAAAACAGATTGAGCTAAAGAAGCCAACCAA
ATCCCACCAAAACCAAGTTGCCACGAGAGTGACCTCTGGCGTCTCACTGCTACACTCCT
GACAGCACCATGACAGTTACAAATGCCATGGCAACATCAGGAAGTTACCCGATATGTCCCA
AAAGGGGGAGGAATGAATAATCCACCCCTGTTAGCAAATAAGCAAGAAATAACCATAAAA
GTGGGCAACCAGCAGCTAGGCGCTGCTTGTATGGAGTAGCCATTCTTTGTTCTT
TACTTCTTAATAAACTTGTCTTACCTTAAAAAA

FIGURE 93

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA54002
><subunit 1 of 1, 544 aa, 1 stop
><MW: 60268, pI: 9.53, NX(S/T): 3
MLLPLLLSSLLGGSQAMDGRFWIRVQESVMVPEGLCISVPCSF SYPRQDWGSTPAYGYWFK
AVTETTKGAPVATNHQSREVEMSTRGRFQLTGDPAKGNCSL VIRDAQM QDESQYFFRVERGS
YVTYNFMNDGFFLKVTVLSFTPQDHNTDLTCHVDFSRKGVSAQRTVRLRVAYAPRDLVIS
ISRDNTPALEPQPQGNVPYLEAQKGQFLRLCAADSQPPATLSWVLQNRVLSSHPWGPRL
GLELPGVKAGDSGRYTCRAENRLGSQQRALDLSVQYPPENLRVMV SQANRTVLENLGNGTSL
PVLEGQSLCLVCVTHSSPPARLSWTQRGQVLSPSQPSDPGVLELPRVQVEHEGEFTCHARHP
LGSQHVSLSLSVHYKKGLISTAFSNGAFLGIGITALLFLCLALIIMKILPKRRTQTETPRPR
FSRHSTILDYINVVPTAGPLAQKRNQKATPNSPRTPPPGAPSPESKKNQKKQYQLPSFPEP
KSSTQAPESQESQEELHYATLNFPGVRPRPEARMPKGTQADYAEVKFQ

Important features:

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 399-418

N-glycosylation site.

amino acids 100-103, 297-300 and 306-309

Immunoglobulins and major histocompatibility complex proteins signature.

amino acids 365-371

FIGURE 94

TGAAGAGTAATAGTTGGAATCAAAAGAGTCAACGCAATGAACTGTTATTACTGCTGCGTTT
TATGTTGGAAATTCTCTCTATGGCCTTGTCTGGAGCAACAGAAAACTCTCAAACAAAGA
AAGTCAAGCAGCCAGTGCGATCTCATTGAGAGTGAAGCGTGGCTGGTGTGGAACCAATT
TTTGTACCAGAGGAAATGAATACGACTAGTCATCACATCGGCCAGCTAACATGTGATTAGA
CAATGGAAACAATTCTTCAGTACAAGCTTGGGAGCTGGAGCTGAAGTACTTTATCA
TTGATGAAAGAACAGGTGACATATGCCATACAGAAGCTGATAGAGAGGAGCGATCCCTC
TACATCTTAAGAGCCCAGGTAATAAGACATCGCACTGGAAGGGCTGGGAACCTGAGTCTGA
GTTTGTCATCAAAGTTCGATTACAATGACAATGAACAAAATCCTAAGATGAACCTTTATG
AGGCCATTGTACCAGAGATGTCTCCAGAAGGAACATTAGTTATCAGGTGACAGCAAGTGAT
GCTGACGATCCCTCAAGTGTTAATATGCTCGCTCTACAGCTTACTCAAGGGCCAGCC
ATATTTTTCTGTTGAACCAACACAGGGAGTCATAAATATCTTCTAAATGGATAGGAAC
TGCAAGATGAGTTTGGTAATCAAGCCAAGGACATGATGGTCAGCCAGGGCGTT
TCTGGACAAACAGGTTATTAATAAACTTTCAGATGTTAATGACAAAAGCCTATATTTAA
AGAAAGTTATACCGCTTGACTGTCTGTGAATCTGCACCCACCTGGACTTCTATAGGAACAA
TCATGGCAATATGATGACATAGGAGGAATGCAGAAATGGATTACGACTTTGAAGAGGG
GATTCGCAACACTTGACATTATTACTAATCAATGAAAGGAATAGTTATTTAA
AAAGAAAGTGGATTTGAGCACCAGACCACCTACGGTATTAGGAAAGTTAAACCCATTC
ATGTTTCTGAGCAGCTCAATGAGTACCACCACCTGAGGTCCACCCACCTTTCAAGATCCAG
GTGGAAAGATGTTGATGGCCTCTTTCCTCTCCATATTAGTTATTTGAAGTTTTGA
AGAAACCCACAGGGATCATTGTAGGTGGGTGTCTGCACAGACCCAGACAAATAGGAAT
CTCCTATCAGGTATTATTACTAGGAGCAAAGTGTCAATATCAATGATAAATGGTACAAT
ACTACAAGTAACTACTGGATCGTGAATCAGTCTGGGTCAACCTAAAGGTATTACGCCAC
AGAAAAATACAAATATAGACAGATCTCTCGATCCACGTTAGTGTCAAGTTCTTAACATCA
ATGATCATGTCTCTGTAGTTCTCTCAATACTAATGTAGACCTTTGTGAAAATGCAGGCTCT
GGTCAGGTTAATCAGGACTACGTGCAGGTGGATAGAGATGAATCCATAGAGACGCACATTT
TTACTTTAAATCTATGTTAAGAGACACTAAATCAAGTTTTACATCAATAGATAAATCAAG
ATAAACACAGGTCTATTTGACTAAAGACTTGGTTAACCTTCAAGAAAGACCTGTCTTC
TACATCTCCATCTTAATGGCCGACAAATGGAAATCCGTCACTACAGATACAAACCCTTAC
CATCCATGTCTGTGACTGGGTGACGATGGGAGCACACAGACCTGCCCAGGTACCCAGGGGTTG
TGCTTTCCATGGATTCAGACAGAAGTTTATCATTGTCTATTCATTTGCATTATGATCATA
TTGGGTTTATTTTTGACTTTGGTTAAACACGGAGAAACAGATTTCTATTTCTGA
GAAAGATGTAAAGATTCAGAGGAATATATCCAAATATGTATGTGAAGGGGTGGAGAAAG
ATACAGAGGCCTTTGAATATAGCAAGGTGGAGGTAGACCCATATGCGGAACGCAAGACAT
CGAAACACACCAAGCGCTGAGATCAGGGCCTATACAGGCAGCTTTGCCAAGTTGGCCCGGA
CAGTGCCCATATCAGGAAATTCATGTGGAAAGGTCGCAAGAAAGGTCAATACTAATGTCCGTGT
CCCCCTCTTTGAATCCCTCCAGACCTACGTCTTTGAGGGAACAGGGTCATTAGGTGGATCC
CTGAGTCTCTTAGAATCAGCAGGTCTGTATCAGGATGAAAAGGTATGATTACCTTAATGAGTTT
GGGACTCGCTTTAAAGATTAGGATGCATGTTTGGTTCTGCAGGTCAAGTCAAATAAATTAG
GCTTTTTACCCATAAATTTAAAGGTCAATGTGTTTCAACCCAAATGGATGTCTAA
AGAGTTTTGTGCCTGGCTATGGGGAAAGCCCTAGGTCTATGGAGTTTTCTGATTCC
CTGGAGTAAAATCTCCATGGTTATTTAAAGCTACACGTCAATGTGTCTATGAACAGAGATGTGG
GGGAAAATGTAAACACATCAGGCTACAGGCATCAAAACCCAGATGTGAAGGTAAAAATAT
TAGGAAGATATAAAGGTAGATGTGAGGAGACACAGATGTAGGTCGATCCTTATGTCGATTAT
CATTATTTACCTTAGGAAAGGTAAAAATACCAACGGAAAAATTTAAAGGTAAAAAAATTT
CAAGTCAAATAGAAATGTACAAATCGAGATAAACTTACATTCATATGAATGTAAAA
ATTGAAAAATGTATAGTCAGAGAAATTTCAGTAAATTTCCATGAAGGTATTTGGCCTTAT
TTAA

FIGURE 95

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53906
><subunit 1 of 1, 772 aa, 1 stop
><MW: 87002, pI: 4.64, NX(S/T): 8
MNCYLLLRFMLGIPLLWPCLGATENSQTKVKQPVRSHLRVKRGWVWNQFFVPEEMNTTSHH
IGQLRSIDLNGNNNSFQYKLLGAGAGSTFIIDERTGDIYAIQKLDREERSLYILRAQVIDIAT
GRAVEPESEFVIKVSDINDNEPKFLDEPYEAIVPEMSPEGTLVIQVTASDADPSSGNNARL
LYSLLQGQPYFSVEPTTGVISSKMDRELQDEYWVIIQAKDMIGQPGALSGTTSVLIKLSD
VNDNKPIFKESLYRLTVSESAPGTTSIGTIMAYDNDIGENAEMDYSIEEDDSQTFDIITNHE
TQEGLIVLKKKVDFEHQNHYGIRAKVKNHHVPEQLMKYHTEASTTFIKIQVEDVDEPPLFLL
PYYVFEVFEETPQGSFVGVVSATDPDNRKSPIRYSITRSKVFNINDNGTITTSNSLDREISA
WYNLSITATEKYNIEQISSIPLYVQVLNINDHAPEFSQYYETYVCENAGSGQVIQTISAVDR
DESIEHHHFYFNLSVEDTNNSSFTIIDNQDNTAVILTNRTGFNLQEEPVFYISILIADNGIP
SLTSTNTLTIHVCDCGDGSQTQTCQYQELVLSMGFKTEVIIAILICIMIIFGFIFLTGLKQ
RRKQILFPEKSEDFRENIFQYDDEGGGEEDTEAFDIAELRSSTIMRERKTRKTTSAEIRSLY
RQSLQVGPDSAIFRKFILEKLEEANTDPCAPPFDLQTYAFEGTGSLAGSLSSLESAVSDQD
ESYDYLNELGPRFKRLACMFGSAVQSNN
```

Important features:

Signal peptide:

amino acids 1-21

Transmembrane domain:

amino acids 597-617

N-glycosylation sites.

amino acids 57-60, 74-77, 419-423, 437-440, 508-511, 515-518,
516-519 and 534-537

Cadherins extracellular repeated domain signature.

amino acids 136-146 and 244-254

FIGURE 96

ATTTCAAGGCCAGCCATATTTNTGTTGAACCAACAACAGGAGTCATAAGAATATTTNTA
AAATGGATAGAGAACTGCAAGATGAGTATTGGGTAATCATTCAAGCCAAGGACATGATTGGT
CAGCCAGGAGCGTTGTNTGGAACAACAAGTGTATTAATTAAACTTCAGATGTTAATGACAA
TAAGCCTATATTTAAAGAAAGTTATACCGCTTGACTGTNTGAATCTGCACCCACTGGGA
NTTNTATAGGAACAATCATGGCATATGATAATGACATAGGAGAGAATGCAGAAATGGATTAC
AGCATTGAAGAGGGATTCGCAAACATTGACATTATT

FIGURE 97

GCAACCTCAGCTCTAGTATCCAGACTCCAGCGCCGCCCGGGCGCGAACCCAAACCCGAC
CCAGAGCTTCTCCAGCGCGCGAGCAGCAGCAGGGCTCCCGCCTTAACCTCCTCCGCGGGG
CCCAGCCACCTTCGGAGTCCGGGTGCCACCTGCAAACCTCTCCGCTTCTGCACCTGCCA
CCCCCTGAGCCAGCGCGGGCCCCGAGCGAGTC**ATGGCCAACGC**GGGGCTGCAGCTGTTGGC
TTCATTCTCGCCTTCCTGGGATGGATCGGCCATCGTCAGCAGCTGCCCTGCCCAAGTGGAG
GATTTACTCCTATGCCGGCGACAAACATCGTGAACGCCAGGCCATGTACGAGGGCTGTGGA
TGTCTCGTGTGCAGAGCACCGGGCAGATCCAGTGCAAAGTCTTGACTCCTGCTGAAT
CTGAGCAGCACATTGCAAGCAACCGTGCCTGATGGTGGCATCCTCCTGGAGTGAT
AGCAATCTTGTGCCACCGTTGGCATGAAGTGTATGAAGTGCTTGGAAAGACGATGAGGTGC
AGAAGATGAGGATGGCTGTCAATTGGGGTGCATATTCTCTTGCAAGGTCTGGCTATT
GTTGCCACAGCATGGTATGGCAATAGAACATCGTTCAAGAATTCTATGACCCATGACCCAGT
CAATGCCAGGTACGAATTGGTCAGGCTCTTCACACTGGCTGGGCTGCTGCTCTCTGCC
TTCTGGGAGGTGCCCTACTTTGCTGTTCTGTCCCCGAAAAACAAACCTCTACCCAACACCA
AGGCCCTATCCAAAACCTGCACCTCCAGCGGGAAAGACTACGTG**TGAC**ACAGAGGAAAAG
GAGAAAATCATGTTGAAACAAACGAAAATGGACATTGAGATACTATCATTAAACATTAGGAC
CTTAGAATTGGTATTGTAATCTGAAGTATGGTATTACAAAACAAACAAACAAAA
ACCCATGTGTTAAACTCAGTGTAAACATGGCTTAATCTTATTTATCTCTTCTCA
ATATAGGAGGGAAAGATTTCCATTGTATTACTGCTTCCATTGAGTAATCATACTCAAAT
GGGGGAAGGGTGCCTTAAATATATAGATATGTATATACATGTTTCTATTAAAAA
ATAGACAGTAAAATACTATTCTCATTATGTTGATACTAGCATACTTAAATATCTCTAAAAT
AGGTAAATGTATTAAATTCCATTGATGAAGATGTTATTGGTATATTCTTTCTCGTCC
TTATATACATATGTAACAGTCAAATATCATTACTCTTCTTCAATTCTCATGCGTGCCTT
CCACAAGACCTAGCCTAATTACCAAGGATGAATTCTTCAATTCTCATGCGTGCCTT
CATATACTTATTTCACCATATAGCACTGCATCGTTATTAGCCCTTAT
TTGTTTGTGTTCAATTGGTCTCTATCTCCTGAATCTAACACATTCTAGCCTACATT
GTTCTAAAGCAAGAAGAATTATTACAAATCAGAACTTGGAGGCAAATCTCTGCATG
ACCAAAGTGATAAATTCTGTTGACCTCCACACAATCCCTGTACTCTGACCCATAGCACT
CTGTTGCTTGAATATTGCAATTGAGTAGCTGCATGCTGTTCCCCAGGTGTT
AACACAACATTATTGATTGAATTGAAAGCTACTTATTGATGTTTATACCCCTAAACT
ACCTTTTGTCCCCATTCTTAATTGTATTGTTCCAAAGTGTAAATTATCATGCGTTTA
TATCTCTTAATAAGGTGGTCTGTTGTAACAAAGTGTAGACTTTCTGGAGTGATA
ATCTGGTGACAAATATTCTCTGTAGCTGTAAAGCAAGTCACCTAACCTTACCTCTT
TTCTATCTGCCAAATTGAGATAATGATACTTAACCAGTTAGAAGAGGTAGTGTGAATTAA
TTAGTTTATATTACTCTTATTGTAACATGAACTATGCCTATGTAGTGTCTTATTGCT
CAGCTGGCTGAGACACTGAAGAAGTCACCTGAACAAACCTACACACGTACCTCATGTGATT
CACTGCCTCCTCTCTACCAAGTCTATTCCACTGAACAAACCTACACACATACCTCAT
GTGGTTCACTGCCTCCTCTACCAAGTCTATTCCACTGAACAAACCTACGCACATAC
CTTCATGTGGCTCAGTGCCTCCTCTACCAAGTCTATTCCACTTCACTGCTGTCT
GACATGTTGTGCTGTTCCATTAAACAATGCTCTTACTTTCCAGTCTGTACAGAATG
CTATTCACCTGAGCAAGATGATGTAATGAAAAGGGTGTGGCACTGGTGTCTGGAGACCTG
GATTGAGTCTGGTGTATCAATCACCGTCTGTGTTGAGCAAGGCAATTGGCTGCTGAA
GCTTATTGCTTCATCTGTAAGCGGGTTGTAATTCTGATCTTCCACCTCACAGTGTGATG
TTGTGGGATCCAGTGAGATAAGATACATGTAAGTGTTGTAATTAAAAGTGTAT
ACTAAGGAAAGAATTGAGGAATTAACTGCATACGTTGGTGTGCTTCAAATGTTGA
AAATAAAAAAAATGTTAAG

FIGURE 98

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52185

><subunit 1 of 1, 211 aa, 1 stop

><MW: 22744, pI: 8.51, NX(S/T) : 1

MANAGLQLLGFI LAFLGWIGAI VSTALPQWRI YSYAGDNIVTAQAMY EGLWMSCVSQSTGQI
QCKVFDSSLNLSSTLQATRALM VVGILLGVIAIFVATVGMKCMK CLEDDEVQKMRMAVIGGA
I FLLAGLAILVATAWYGNRIVQEFYDPMTPVNARYEFGQALFTGWA AASLCLLGGALLCCSC
PRKTTSYPTPRPYPKPAPSSGKDYV

Important features:

Signal peptide:

amino acids 1-21

Transmembrane domains:

amino acids 82-102, 118-142 and 161-187

N-glycosylation site.

amino acids 72-75

PMP-22 / EMP / MP20 family proteins

amino acids 70-111

ABC-2 type transport system integral membrane protein

amino acids 119-133

FIGURE 99

TTCTGGCAAACCCGGGCTNCAGCTGTTGGGCTTCATCTGCCTCCTGGATGGATGGC
GCCATCNTCACACTGCCCTCCCCAGTGGAGGATTTACTCCCTATGCTGGCGACAACATCG
TGACCGCCCAGCCCATGTACGAGGGCTGTGGATGTCCNGCGTGTGCAGAGCACCGGGCAG
ATCCAGTGCAAAGTCTTGACTCCTGCTGAATCTGAGCAGCACATTGCAAGCAACCGTGC
CTTGATGGTGGTTGGCATCCTCCTGGAGTGATAGCAATCTTGTGGCCACCGTTGGCATGA
AGTGTATGAAGTGCTTGGAAAGACGATGAGGTGCAGAAGATGAGGATGGCTGTCAATTGGGGC
GCGATATTCTTCTTGCAGGTCTGGCTATTAGTTGCCACAGCATGGTATGGCAATAGAAN
CNTTCAACANTTCTATGACCCATGACCCAGTCAATGCCAGGTACGAATTGGTCA
GGCTCTCTTCACTGGCTGGCTGCTGCTCTCTGCCTCTGGGAGGTGCCCTACTTGCT
GTTCCTGTCCC

FIGURE 100

ACCCTTGACCAACGCGGCCCCCGACCGNTTCATGGCCAAACGCGGNCTCCAGCTGTTGG
GCTTCATTCTCCCTTCCTGGGATGGACCGGGCGCCATCNTCAGCACTGCCCTGCCCTAGTG
GAGGATTACTCCTATNCCGGCNACAAACATCGTACCGCCCAGGCCNTGTACGAGGGGCTGT
GGATGTCCTGCGTGTGCAGAGCACCGGGCAGATCCAGTGCAAAGTCTTGACTCCCTGCT
GAATCTGAGCAGCACATTGCAAGCAACCCGTGCCTTGATGGTGGTTGGCATCCTCCTGGGAG
TGATAGCAATCTTNNNTGCCACCGTTGTNNNTGAAGTGTATGAAGTGCTTGGAAAGACGATGA
GGTGCAGAAGATGAGGATGGCTGTCATTGGGGCGCGATATTCTTCTTGCAGGTCTGGCTA
TTTAGTTGCCACAGCATGGTATGGCAATAGAATCGTTCAAGAATTCTATGACCCTATGACCGA

FIGURE 101

GGGCCGACCATTATCCAACCGGGNTCACTGTTGGCTCATCTCCCTCCTGGATGAANCGCGC
CATCNTCAGACTCCCTGCCCATGGAGATTNNCCTATGCTGGCGACAAACATCNTGACCCCC
AGCCATGTACGAGGGGTTGAACGTCNGCGTGTGCAGANCACCGGGCAGATCCAGTGC
AGTCTTGACTCCTGCTGAATCTGNGCAGCACATTGCAGCAACCCNTGCCCTGATGGTGGT
TGGCATCCTCCTGGGAGTGATAGCAATCTTGTGGCCACCCTGGCATGAAGTGTATGAAGT
GCTTGGAAAGACGATGAGGTGCAGAAAGATGAGGATGGCTGTCATTGGGGCGCGATATTCTT
CTTGCAGGTCTGGCTATTNNNNTGCCACAGCATGGTATGGCAATAGAATCGTTCAAGAAT
TCTATGACCCCTATGACCCAGTCAATGCCAGGTACGAATTGGTCAGGCTCTTCACTGGC
TGGGCTGCTGTTCTCTGCCTCTGGAGGTGCCCTACTTGCTGTTCTGC
GA

FIGURE 102

ATTCTCCCTCCTGGATGGATCGCNCCACCGTCACATTGCCTCCCCANTGGAGGATTNAC
TCCTATGCTGGCGACAACATCGTACCCCCCAGGCCATTACCGAGGGGCTTGGATGTCNT
GCNTGTCGCAGAGCACCGGGCAGATCCCAGTGCAAAGTCTTGACTCCTGCTGAATCTGAG
CAGCACATTGCAAGCAACCGTGCCTGATGGGTTGGCATCCTCCTGGAGTGATAGCAAC
CTTGTCGCCACCGTTGGCATGAAGTGTATGAAGTGCTTGGAAAGACGATGAGGTGCCAGAAG
ATGAGGATGGCTGTCATTGGGGCGCGATATTCTTGTGCAGGTCTGGCTATTTAGTNGC
CACAGCATGGTATGGCAATAGANTNNTCNNGNNNTCTATGACCCATGACCCAGTCAATG
CCAGGTACGAATTGGTCAGGCTCTTCACTGGCTGGCTGCTGCTCTCTGCCTTCTG
GGAGGTGCCCTACTTGCTGTTCTGTCCC

FIGURE 103

AGAGCACCGGCAGATCCCAGTNCAAAGTCTTGACCCTGCTGAATCTGAGCAGCACATTNC
AAGCAACCCCTGCCTTGAAGGTGGTTGNCATCCCCCTGGGAGTGAATAGCAATCTTGTG
GCCACCCTGGCATGAAGTNTATGAAGTGCTTGAAGACGATGAGGTGCAGAAGATGAGGAT
GGCTGTCATTGGGGCGCGATATTCTTGCAGGTCTGGCTATTTAGTNCCACAGCAT
GGTATGGCAATAGNATNNNTCGNGGNTTCTATGACCTATGACCCAGTCAATGCCAGGTAC
GAATTGGTCAGGCTCTCTCACTGGCTGGCTGCTGTTCTCTGCCTCTGGGAGGTGC
CCTACTTTGCTGTTCTGTCCCCGAA

FIGURE 104

AGCAATGCCCTGCCCCCAGTGGAGGATTAATTCTATGNTGGGGACAACATTGTGACNGCCC
AGGCCATGTACGGGGGGCTGTGGATGTCCTGCGTGTGCAGAGCACCGGGCAGATCCAGTGC
AAAGTNTTGACTCCTGCTGAATTGAGCAGCACATTGCAAGCAACCGTGCCTGATGGT
GGTGGCATCTCCTGGAGTGATAGCAATCTTGTGCCACCGTGGNAATGAAGTGTATGA
AGTGCTTGGAAAGACGATGAGGTGCAGAAGATGAGGATGGCTGTCATTGGGGCGCGATATT
CTTNTTGCAAGGTCTGGCTATTTAGTTGCCACAGCATGGTATGGCAATAGAATNGTTCAAGA
ATTTATGACCCATGACCCAGTCAATGCCAGGTACGAATTGGTCAGGCTTNTTCACTG
GCTGGGCTGCTGCTNTTCTGCCTNTGGGAGGTGCCCTANTTGCTGTTCCCTGCGAAC

FIGURE 105

TCATAGGGGGCGCGATATTTTCTTGCAGGTNTGGTTATTTAGTTGCCACAGCATGGTA
TGGCAATAGAATCGTTCAAGAATTNTATGACCCTATGACCCAGTCAATGCCAGGTACGAAT
TTGGTCAGGCTCTNTCACTGGNTGGGCTGCTGCTCTNTNNGCCTNTGGGAGGTGCCCTA
CTTGCTGTTCTG

FIGURE 106

TTCTGGGATGGATCCGCCCATCNTCACATGCCCTGCCCNNTGGAGAGTTACNCCTATGC
TGGCGAACAAACATCNTGACCGCCCAGGCCATGTACGAGGGCTGTGGAATGTCTGCGTGTC
CCAGAGCACCAGGGCAGATCCAGTGCAAAGTCTTGACTCCTGCTGAATCTGAGCAGCACAT
TGCAAGCAACCNTGCCTTGATGGTGGTTGGCATCCTCCTGGAGTGATAGCAATCTTGTGG
CCACCGTTGGCATGAAAGTGTATGAAGTGCTTGGAAAGACGATGAGGTGCAGAAGATGAGGAT
GGCTGTCATTGGGGCGCGATATTCTTCTGCAGGTCTGGCTATTTAGNNGCCACAGCAT
GGTATGGCAATCAGACCCNNTCANAAACTCTATGACCCCTATGACCCAGTCAATGCCAGGTA
CGAATTTGGTCAGGCTCTTCACTGGCTGGCTGCTGCTTCTCTGCCTTCTGGAGGTG
CCCTACTTTGCTGTTCTGTCCCCGAAAAACACCTTTACCCACG

FIGURE 107

CGGGGCTGCAGCTGTTGGGCTTCATCTCGCTTCCTGGGATGGAATCGCGCCATCGTCAGCA
CTGCCCTGCCCATGGAGGATTTACTCNTATGCTGGCGACAACATCGTACCCAGGCCA
TGTACGAGGGCTGTGGATGTCNGCGTGTGCAGAGCACCAGATCCAGTGCAAAGTCT
TTGACTCCTGCTGAATCTGAGCAGCACATTGCAAGCAACCNTGCCTGATGGTGGTTGGCA
TCCTCCTGGGAGTGATAGCAATCTTGTGCCACCGTTGGCATGAAGTGTATGAAGTGCTTG
GAAGACGATGAGGTGCAGAAGATGAGGATGGCTGTCAATTGGGGCGCGATATTCTTCTGC
AGGTCTGGCTATTNTAGTTGCCACAGCATGGTATGCAATAGAATCGTTCAAGAATTCTAT
GACCCTATGACCCAGTCAATGCCAGGTACGAATTGGTCAGGCTCTTCACTGGCTGGC
TGCTGCTTCTCTGCCTCTGGGAGGTGCCCTACTTGCTGTTCTGCGAA

FIGURE 108

GGGTGCCGTCAGCTGCCGGCACCGCGGCCCTGCCCTGCCCTCCGCCCTGCGCTGCAC
CGCGTAGACCGACCCCCCCCCTCCAGCGGCCACCCGGTAGAGGACCCCCGCCGTGCCCG
ACCGGTCCCCGCCCTTTGTAAAACCTAAAGCGGGCGCAGCATTACGCTCCGCCCGT
GACCTCTCAGGGTCTCCCGCAAAGGTGCTCCGCCGCTAAGGAAC**ATGGCGAAGGTGGAG**
CAGGTCTGAGCCTCGAGCCGAGCACGAGCTCAAATTCCGAGGTCCCTCACCGATGTTGT
CACCACCAACCTAAAGCTTGGCAACCCGACAGACCGAAATGTGTGTTAAGGTGAAGACTA
CAGCACCAACGTAGGTACTGTGTGAGGCCAACAGCGGAATCATCGATGCAGGGCCTCAATT
AATGTATCTGTGATGTTACAGCCTTCGATTATGATCCAATGAGAAAAGTAAACACAAGTT
TATGGTTCAGTCTATGTTGCTCCAAC TGACACTTCAGATATGGAAGCAGTATGGAAGGGAGG
CAAAACCGGAAGACCTTATGGATTCAAACCTAGATGTGTGTTGAATTGCCAGCAGAGAAT
GATAAACACATGATGTAGAAATAAATAAATTATATCCACAACACTGCATCAAAGACAGAAAC
ACCAATAGTGTCTAAGTCTTGAGTTCTTGGATGACACCGAAGTTAAGAAGGTTATGG
AAGAATGTAAGAGGCTGCAAGGTGAAGTTAGAGGCTACGGGAGGAACAAGCAGTTCAAG
GAAGAAGATGGACTCGGGATGAGGAAGACAGTGCAGAGCAACAGCCCCATTCAGCATTAGC
CCCAACTGGGAAGGAAGAAGGCCCTAGCACCCGGCTTGGCTCTGGTGGTTTGTCTTTA
TCGTTGGTGTATTGGGAAGATTGCCTTGT**AGAGGTAGCATGCACAGGATGGTAAATTG**
GATTGGTGGATCCACCATATCATGGATTAAATTATCATAACCAGTGTAAAAAGAAATT
AATGTATGATGACATCTCACAGGTCTTGCCTTAAATTACCCCTCCCTGCACACACATACAC
AGATAACACACACAAATATAATGTAACGATCTTTAGAAAGTTAAAATGTATAGTAACTG
ATTGAGGGGGAAAAAGAATGATCTTATTAAATGACAAGGGAAACCATGAGTAATGCCACAAT
GGCATATTGTAAATGTCATTTAAACATTGGTAGGCCCTGGTACATGATGCTGGATTACCTC
TCTTAAAATGACACCCCTTCCTGCCCTGGTGTGGCTGCCCTGGGAGCTGGAGGCCAGCAT
GCTGGGGAGTGCAGCTCCACACAGTAGTCCCCACGTGCCACTCCGCCAGGCTG
CTTCCGTCTCAGTCTGTCCAAGCCATCAGCTCCTGGACTGATGAAACAGACTCAGA
AGCCAAAGGAATTGCACTGTGGCAGCAGTACACTCGTCATAAGTGAGAGGCGTGTGT
TGACTGATTGACCCAGCGCTTGGAAATAATGGCAGTGCTTGTACTTAAAGGGACCAA
GCTAAATTGTATTGGTCATGTAGTGAAGTCAAACTGTTATTCAAGAGATGTTAATGCATA
TTAACCTATTAATGTATTCATCTCATGTTTCTTATTGTCACAAGAGTACAGTTAATGC
TGCCTGCTGCTGAACCTCTGTTGGGTGAACGGTATTGCTGGAGGGCTGGCTCCTCT
GTCTCTGGAGAGTCTGGTATGTGGAGGTGGGTTATTGGATGCTGGAGAAAGAGCTGCCA
GGAAGTGTTTCTGGTCAGTAAATAACAACGTACAGGGAGGGAAATTCTCAGTAGTG
ACAGTCACACTAGGTTACCTTTAATGAAGAGTAGTCAGTCTTAGATTGTTCTTATA
CCACCTCTCAACCATTACTCACACTCCAGCGCCCAGGTCCAAGTCTGAGCCTGACCTCCCC
TTGGGGACCTAGCCTGGAGTCAGGACAATGGATGGCTGAGAGGGTTAGAAGCGAGGGC
ACCAAGCAGTTGTGGGTGGGAGCAAGGGAAAGAGAAAAGTCTCAGCGAATCCTCTAGTAC
TAGTTGAGAGTTGACTGTGAATTAAATTGTCATGCAAAAGACCAACCCAGTTCTGTTGA
CTATGTAGCATCTTGAAGAAAAATTATAATAAGCCCCAAAATTAAAGAAAA

FIGURE 109

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53977
<subunit 1 of 1, 243 aa, 1 stop
<MW: 27228, pI: 7.43, NX(S/T): 2
MAKVEQVLSLEPQHELKFRGPFTDVTTNLKGNPTRNVCFKVKTAPRRYCVRPNSGIID
AGASINVSVMLQPFDYDPNEKSCHKFMVQSMFAPTDTSMDMEAVWKEAKPEDLMDSKLRCVFE
LPAENDKPHDVEINKIISTTAKTETPIVSKSLSSLDTEVKVMECKRLQGEVQRLREE
NKQFKEEDGLRMRKTVQNSPISALAPTGKEEGLSTRLLALVVLFFIVGVIIGKIAL
```

Important features:

Transmembrane domain:

amino acids 224-239

N-glycosylation site.

amino acids 68-71

N-myristoylation site.

amino acids 59-64, 64-69 and 235-240

FIGURE 110

GTCAGTCTTAGATTGTCCTTATCCCACCTTCAACCANTACTCACATTCNAGGCCAG
GTCCANGTCTGAGCCTGACTTCCCCTGGGGACCTAGCCTGGAGTCAGGACAATGGNTGGG
CTGCAGAGGNTTAGAAGCGAGGGCACCAGCAGTTGGTGGGAGCAAGGGNNGAGAGAAA
CTCTTCAGCGAATCCTCTAGTACTAGTTGAGAGTTGACTGTGAATTAATTATGCCATA
AAAGACNAACCCAGTTCTGTTGACTATGTAGCATCTGAAAAGAAAATTATAATAAGCC
CCAAAATTAAGAATTCTTGTCACTTGCTCATGGGGGAATTATTATTTT
ATCATTTTATTATTTGCCATTGGAAGGTTAACTTAAAATGAGC

FIGURE 111

TATTGTAAAGGCCATTTAAACCATTGGTAGGCCTTGGTACATGATGCTGGATTACCTCCTT
AAATGACACCNTTCCTCGCCTGTTGGTCTGGCCNTGGGGAGCTGGAGCCCCAGCATGCTG
GGGAGTGCAGTCAGCTCCACACAGTAGTCCCCACGTGGCCCACCTCCGGCCCAGGCTGCTT
CCGTGTCTTCAGTTCTGTCCAAGCCATCAGCTCCTGGACTGATGAACAGAGTCAGAACGCC
CAAAGGAATTGCCACTGTGGCAGCAGCATCAGACGTACTCGTCATAAGTGAGAGGCGTGTGTTGA
CTGATTGACCCAGCGCTTGGAAATAATGGCAGTGCTTGTTCACTTAAAGGGACCAAGCT
AAATTGTATTGGTTCATGTAGTGAAGTCAAACGTATTCAAGAGATGTTAATGCATATTA
ACTTATTTAATGTATTCATCTCATGTTTCTTATTGTACAAAGAGTACAGTTAATGCTGCG
TGCTGCTGAACCTGTTGGGTGAACGGTATTGCTGCTGGAGGGCTG

FIGURE 112

CCCTGGTGGTTTGTCTTAATTCGTTGGTGTATTNTTGGGAAGATTGCTGTAGAGGTA
GNATGCACCNGGCTGGTAAATTGGATTGGTGGATCCACCATATCCATGGGATTAAATTAT
CATAACCATGTGTAAAAAGAAATTAATGTATGATGACATNTCACAGGTATTGCCTTAAATT
ACCCATCCCTGNANACACATACACAGATACACANANACAAATNTAATGTAACGATNTTTAG
AAAGTTAAAAATGTATAGTAAC

FIGURE 113

GGTGGCCCATTCCCGGCCAGGCTGCTTCCGGTNTTCAGTTCTGTCCAAGCCATCAGCTCC
TTGGGACTGATGAACAGAGTCAGAACGCCAAAGGAATTGCACTGTGGCAGCATNAGACGTAC
TTGTNATAAGTGAGAGGCCTGTGTTGACTGATTGACCCAGCGCTTGAAATAATGGCAGT
GCTTGTTCACTAAAGGGACCAAGCTAAATTGTATTGGTCATGTAGTGAAGTCAAACGT
TTATTCAGAGATGTTAATGCATATTAANTTATTAATGTATTTNATNTCATGTTTCTTA
TTGTCACAAGAGTACAGTTAATGCTCGTGCTGAANTNTGTTGGGTGAACGTGGTATTGC
TGCTGGAGGGCTGTGGGCTCCTCTGTCTTGAGAGTCTGGTCATGTGGAGGTGG

FIGURE 114

TGCTTCCGTCTCAGTTCTGTCCAAGCCATCAGCTCCTGGGACTTGATGAACAGAGTC
AGAAGCCCAAAGGAATTGCACTGTGGCAGCATCAGACGTACTCGTCATAAGTGAGAGGCGTG
TGTTGACTGATTGACCCAGCGCTTGGAAATAATGGCAGTGCTTGTTCACTTAAAGGGAC
CAAGCTAAATTGTATTGGTTCATGTAGTGAAGTCAAACTGTTATTCAAGAGATGTTAATGC
ATATTAACTTATTTAATGTATTCATCTCATGTTTCTTATTGTCAACAAGAGTACAGTTAA
TGCTGCGTGC

卷之三

FIGURE 115

AAACCTTAAAAGTTGAGGGAAAAGAATGATCCTTATTAATGACAAGGGAAACCNTGN
GTTAATGCCACAATGGCATATTGTAAATGTCATTTAACATTGGTAGGCCTGGTACATGATGC
TGGATTACCTCTCTTAAATGACACCCCTCCTGCCTGTTGGTGTGCCCTGGGAGCTN
GAGCCCAGCAGTGGGAGTGCAGCTGCTCCACACAGTAGTCCCCANGTGGCCANTCCC
GGCCCAGGCTGCTTCCGTCTTCAGTTCTGTCCAAGCCATCAGCTCCTGGGANTGATGA
ACAGAGTCAGAAGCCAAAGGAATTGCANTGTGGCAGCATCAGANGTANTNGTCATAAGTGA
GAGGCAGTGTGTTGANTGATTGACCCAGCGCTTGGAAATAATGGCAGTGCTTGTTGTT
AAAGGGNCCAAGNTAAATTGTATTGGTTCATGTAGTGAAGTCAAANTGTTATTCA
GAGAGATGTTAATGCATATTTAANTTATTAAATGTATTCATNTCATGTTTCTTATTGTCACAAGGGT
ACAGTTAATGCTGCGTGCTGCTGAANTCTGTTGGGTGAANTGGTATTGCTG

FIGURE 116

GGCCCTTGGGAGCTGGAGCCCAGCATGCTGGGAGTGCAGCTCCACACAGTAGTCCC
CACGTGGCCCCTCCGGCCAGGCTGCTTCCGTCTCAGTTCTGTCCAAGCCATCAGC
TCCTTGGACTGATGAACAGAGTCAGAAGCCAAAGGAATTGCACTGTGGCAGCAGACG
TACTCGTCATAAGTGAGAGGGCGTGTGTTGACTGATTGACCCAGCGCTTGGAAATAATGGC
AGTGCTTGTTCACTTAAAGGGACCAAGCTAAATTTGTATTGGTTCATGTAGTGAAGTC
CTGTTATTCAAGAGATGTTAATGCATATTTAACTTAAATGTATTTCATCTCATGTTTC
TTATTGTCACAAGAGTACAGTTAATGCTGCGTGCTGCTGAACCTCTGTGGTGAACGGTAT
TGCTGCTGGAGGGCTGTGGGCTCCTCTGTCTGGAGAGTCTGGTCATGTGGAGGTGG

FIGURE 117

GGAGCTCCGGTCTGGCCGGCCTGGCGGGCGGCCTCGGCTCAGGCTGGCTGAGA
GGCTCCCAGCTGCAGCGTCCCCGCCCTCCTCGGGAGCTCTGATCTCAGCTGACAGTGCC
CTCGGGGACAAACAAGCCTGGCAGGGTCTCACTTTGTTGCCAGGCTGGAGTTAGTCCA
TGATCATGGTTACTGCAGCCTGACCTCCTGGTTCAAGCGATCTGCTGAGTAGCTGG
CTACAGGACAAAATTAGAAGATCAAAATGGAAAATATGCTGCTTGGTGATATTTTCA
CCTGGGTGGACCCTCATTGATGGATCTGAAATGGAATGGGATTTATGTGGCAGTGAGAAA
GGTACCCCGATTGTCAGTGAAAGGACTTCCATCTCACCGCCCGCATTGAGGCAGATG
CTAAGATGATGGTAAATACAGTGTGGCATCGAATGCCAGAAAGAACTCCAACTCCAGC
CTTCTGAATTGGAGGATTATCTTCATGAGACTGTCTTGAGAATGGCACCCGAACCTT
AACCAAGGGTCAAAGATTGGTCTGAGCCACTCAAAATATCACCAAAAGGGAG
TATCTGTTAGGAGAAAGAGACAGGTGTATGGCACCGACAGCAGGTCAGCATCTGGACAAA
AGGTTCTAACCAATTCCCTTCAGCACAGCTGTGAAGCTTCCACGGGCTGTAGTGGCAT
TCTCATTCCCCTCAGCATGTTCTAAGTGTGCCCAGTGTGTTCATGATGGAAGGACTATG
TCAAAGGGAGTAAAAGCTAACGGTAGGGTGTGAAGATGAGGAATAAAAGTGGAGGAAG
AACGTCGAGGTTCTAACAGGGAGCAGGAGAGAAGCTAGTGGTGGTACCAAGAGAGGGTAC
CAGAGAGCATCTGCAGGAGAGCGAACGGTGGGAGAAGAAGAAAAAAATCTGGCCGGGTC
AGAGGATTGCCAACGGGCTTCCTTCAGTGGACCCGGTCAAGAATACCCACATTCCG
AAGGGCTGGGACGGAGGAGCATGGGGACGCTACCTTGGACTATGACTATGCTCTTCTGGA
GCTGAAGCGTGCTCACAAAAGAAATACATGGAACCTGGATCAGCCAAAGATCAAGAAA
TGCCTGGTGAATGATCCACTTCTCAGGATTGATAACGATAGGGCTGATCAGTTGGTCTAT
CGGTTTGCACTGTGCTCGACGAATCCAATGATCTCCTTACCAATACTGCGATGCTGAGTC
GGGCTCACCAGTTGGGCTATCTGCGTCTGAAAGATCCAGACAAAAGAATTGGAAGC
GCAAAATCATTGCGGTCTACTCAGGGCACCAGTGGTGGATGTCACGGGTTCAAGGGAC
TACAACGTTGCTGTCGACACTCCCCTAAAATACGCCAGATTGCCCTCTGGATTACGG
GAACGATGCCAATTGTGCTAACGG**TAA**CAGAGACCTGAAACAGGGCGGTATCATCTAAA
TCACAGAGAAAACCAGCTCGCTTACCGTAGTGAGATCACTCATAGTTATGCCTGGACTT
GAACCTGTCAATAGCATTCAACATTTCAAAATCAGGAGATTTCGTCCATTAAAAAA
TGTATAGGTGCAGATATTGAAACTAGGTGGCACTTCATGCCAAGTATATACTCTTCTTA
CATGGTGTGAGTTCTATTGTAGAAAAATTGTTGCCCTTAAAGACACACTT
AACACCTCAAACAGGTATTATAAAACATGTGACTCCTTAATGGACTTATTCTCAGGGTCC
TACTCTAACAGAAATCTAACAGGATGCTGGTTGTATTAAATGTGAAATTGCATAGATA
GGTAGATGGTAAAGCAATTAGTATCAGAATAGAGACAGAAAGTTACAACACAGTTGTACTA
CTCTGAGATGGATCCATTCACTGCTCATGCCCTCAATGTTATATTGTGTTATCTGTTGGGTCT
GGGACATTAGTTAGTTAGTTAGAAGAATTACAAATCAGAAGAAAAGCAAGCATTATA
CAAAACTAATAACTGTTACTGCTTAAGAAATAACAATTACAATGTTAGTATTATTTAAAA
TGGGAGAAATAGTTGTTCTATGAAATAACCTAGTTAGAAATAGGGAAGCTGAGACATT
TAAGATCTCAAGTTTATTAACTAATACCTAAATATGGACTTTCATGTATGCATAGGG
AAGACACTCACAAATTATGAATGATCATGTGTTGAAAGCCACATTATTATGCTATACAT
TCTATGTATGAGGTGCTACATTAGGACAAAGAATTCTGTAATCTTTCAAGAAAGAGT
CTTTCTCCTGACAAATCCAGTTGTATGAGGACTATAGGGTGAATTCTCTGATTAG
TAATTAGATATGTCCTTCCTAAAGAATAAAATTATGAATATGA

FIGURE 118

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57253
<subunit 1 of 1, 413 aa, 1 stop
<MW: 47070, pi: 9.92, NX(S/T): 3
MENMLLWLIFFTPGWTLIDGSEMEWDFMWHLRKVPRIVSERTFHLTSPAFEADAKMMVNTVC
GIECQKELPTPSLSELEDYLSYETVFENGTRTLTRVKVQDLVLEPTQNIITKGVSRRKRQV
YGTDSRFSILDKRFLTNFPFSTAVKLSTGCGILISPQHVLTAAHCVHDGKDYVKGSKKLRV
GLLKMNRNKGKRRGSKRSRREASGGDQREGTREHLQERAKGGRRKSGRQRIAEGRPS
FQWTRVKNTHIPKGWARGGMGATLDYDYALLELKRAHKKKYMELGISPTIKKMPGGMIHFS
GFDNDRADQLVYRFCVSDESNDLLYQYCDAESGSTGSGVYRLKDPLKKNWRKIIAVYSG
HQWVDVHGVQKDYNVAVRITPLKYAQICLWIHGNDANCAYG
```

Important features:

Signal peptide:

amino acids 1-16

N-glycosylation sites.

amino acids 90-93, 110-113 and 193-196

Glycosaminoglycan attachment site.

amino acids 236-239

Serine proteases, trypsin family, histidine active site.

amino acids 165-170

FIGURE 119

AATGTGAGAGGGGCTGATGGAAGCTGATAGGCAGGACTGGAGTGTAGCACCAGTACTGGAT
GTGACAGCAGGCAGAGGAGCACTTAGCAGCTTATTCACTGTCGATTCTGATTCCGGCAAGG
ATCCAAGC**ATGGA**ATGCTGCCGTGGCAACTCCTGGCACACTGCTCCTCTTCTGGCTTC
CTGCTCCTGAGTTCCAGGACCGCACGCTCCGAGGAGGACGGACGGCTATGGATGCCTG
GGGCCATGGAGTGAATGCTCACGCACCTGCAGGGAGGGGCTCCTACTCTGAGGCCT
GCCTGAGCAGCAAGAGCTGTGAAGGAAGAAATATCCGATAACAGAACATGCAGTAATGTGGAC
TGCCCACCAGAAGCAGGTGATTCCGAGCTCAGCAATGCTCAGCTATAATGATGTCAAGCA
CCATGGCCAGTTTATGAATGGCTCCTGTCTAATGACCTGACAACCCATGTTCACTCA
AGTCCAAGCAAAGGAACAAACCTGGTTGTAAGTAGCACCTAAGGTCTTAGATGGTACG
CGTTGCTATAAGAAATCTTGGATATGTGCATCAGTGGTTATGCCAAATTGTTGGCTGCGA
TCACCAGCTGGGAAGCACCGTCAAGGAAGATAACTGTGGGTCTGCAACGGAGATGGTCCA
CCTGCCGGCTGGTCCGAGGGCAGTATAAATCCCAGCTCTCCGCAACCAAATCGGATGATACT
GTGGTTGCACTTCCCTATGGAAGTAGACATATTGCCCTGTCTTAAAGGTCTGATCACTT
ATATCTGAAACCAAACCTCCAGGGACTAAAGGTGAAACAGTCTCAGCTCCACAGGAA
CTTCCTTGTGGACAATTCTAGTGTGGACTTCCAGAAATTCCAGACAAAGAGATACTGAGA
ATGGCTGGACCACTCACAGCAGATTTCATTGTCAAGATTGTAACCTGGGCTCCGCTGACAG
TACAGTCCAGTTCATCTTCTATCAACCCATCATCCACCGATGGAGGGAGACGGATTCTTTC
CTTGCTCAGCAACCTGTGGAGGAGTTATCAGCTGACATCGGCTGAGTGCTACGATCTGAGG
AGCAACCGTGTGGTTGCTGACCAATACTGTCACTATTACCCAGAGAACATCAAACCAAACC
CAAGCTTCAGGAGTGCAACTGGATCCTGTCCAGCCAGTGACGGATAACAGCAGATCATGC
CTATGACCTCTACCATCCCTTCCTCGGTGGAGGCCACCCATGGACCGCGTGTCTCC
TCGTGTGGGGGGGCATCCAGAGCCGGCAGTTCTGTGTGGAGGAGGACATCCAGGGCA
TGTCACTTCAGTGGAAAGAGTGGAAATGCATGTACACCCCTAAAGATGCCATCGCGCAGCCCT
GCAACATTTTGACTGCCCTAAATGGCTGGCACAGGAGTGGTCTCCGTGCACAGTGACATGT
GCCAGGGCCTCAGATAACCGTGTGGCCTCTGCATCGACCATCGAGGAATGCACACAGGAGG
CTGTAGCCAAAAACAAAGCCCCACATAAAAGAGGAATGCATCGTACCCACTCCCTGCTATA
AACCCAAAGAGAAACTCCAGTCGAGGCCAAGTTGCCATGGTTCAAACAAGCTCAAGAGCTA
GAAGAAGGAGCTGCTGTCAAGAGGCCCTCG**TAA**AGTTGAAAGCACAGACTGTTCTATA
TTGAAACTGTTGTTAAAGAAAGCAGTGTCTCACTGGTTGAGCTTCATGGTTCTGA
ACTAAGTGTAAATCATCTCACCAAGCTTTGGCTCTCAAATTAAAGATTGATTAGTTCAA
AAAAAAA

FIGURE 120

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58847
<subunit 1 of 1, 525 aa, 1 stop
<MW: 58416, pi: 6.62, NX(S/T): 1
MECCRATPGTLLLFLAFLSSRTARSEEDRGLWDAWGPWSECSRTCGGGASYSLRRCLS
SKSCEGRNIRYRTCSNVDCPPEAGDFRAQQCSAHNDVKHHGQFYEWLPVSNDPDNPCSLKCQ
AKGTTLVELAPVLDGTRCYTESLDMCISGLCQIVGCDHQLGSTVKEDNCGVCNGDGSTCR
LVRGQYKSQLSATKSDDTVVALPYGSRHIRLVLKGPDHLYLETKTLQGTKGENSLSSGTFL
VDNSSVDFQKFPDKEILRMAGLTADFIVKIRNSGSADSTVQFIFYQPIIHRWRETDFFPCS
ATCGGGYQLTSAECYDLRSNRVVADQYCHYYPENIKPKPKLQECNLDPCPASDGYKQIMPYD
LYHPLPRWEATPWTACSSSCGGGIQSRAVSCVEEDIQGHVTSVEEWKCMYTPKMPIAQPCNI
FDCPKWLAQEWSPECTVTCGQGLRYRVVLCIDHRGMHTGGCSPTKPHIKEECIVPTPCYKPK
EKL PVEAKLPWFQQAQELEEGAAVSEEPS
```

Important features:

Signal peptide:

amino acids 1-25

N-glycosylation site.

amino acids 251-254

Thrombospondin 1

amino acids 385-399

von Willebrand factor type C domain proteins

amino acids 385-399, 445-459 and 42-56

FIGURE 121

CGGACGCGTGGCGGCGCTCGGAACTCCCGTGGAGGGCCGGTGGGCCCTCGGCCTGAC
AGATGGCAGTGGCCACTGCGCGGCAGTACTGGCCGCTCTGGCGGGCGCTGTGGCTGGCG
GCCCGCCGGTCTGTGGGCCAGGGTCCAGCGCTGCGCAGAGGCGGGACCCGGCCTCAT
GCACGGGAAGACTGTGCTGATCACCGGGCGAACAGCGGCCTGGCCGCCACGGCCCG
AGCTACTGCGCTGGAGCGCGGGTGTATGGCTGCCGGACCGCGCGCGCCGAGGAG
GCGCGGGTCAGCTCCGCCGAGCTCCGCCAGGCCGGAGTGCAGGCCAGAGCCTGGCG
CAGCGGGGTGGCGAGCTCATAGTCGGAGCTGGACCTCGCTCGCTCGCTCGGTGCG
CCTTCTGCCAGGAAATGCTCCAGGAAGAGCCTAGGCTGGATGTCTGATCAATAACGAGGG
ATCTTCCAGTGCCCTTACATGAAGACTGAAGATGGTTGAGATGCAGTCGGAGTGAACCA
TCTGGGCACTTCTACTCACCAATCTCCTGGACTCCTCAAAAGTTCAAGCTCCAGCA
GGATTGTGGTAGTTCTTCAAACCTTATAAAACGGAGACATCAATTGATGACTTGAAC
AGTGAACAAAGCTATAATAAAAGCTTTGTTAGCCGGAGCAAACGGCTAACATTCTTT
TACCAAGGAACCTAGCCGCCGCTTACAAGGCACAAATGTCACCGTCAATGTGTTGCATCCTG
GTATTGTACGGACAAATCTGGGAGGCACATACACATTCCACTGTTGGTCAAACCAACTCTTC
AATTTGGTGTATGGCTTTTCAAACCTCCAGTAGAAGGTGCCAGACTTCCATTATT
GGCCTCTCACCTGAGGTAGAAGGAGTGTCAAGGAAAGATACTTGGGATTGTAAAGAGGAAG
AACTGTTGCCAAAGCTATGGATGAATCTGTTGCAAGAAAACCTGGGATATCAGTGAAGTG
ATGGTTGGCCTGCTAAAGGAACAAGGAGTAAAGAGCTGTTATAAAACTGCATATCAG
TTATATCTGTGATCAGGAATGGTGTGGATTGAGAACTTGTACTTGAAGAAAAAGAATTG
ATATTGGAATAGCCTGCTAACAGAGGTACATGTTGGTATTGGAGTTACTGAAAAATTATT
TGGGATAAGAGAATTCAAGGAAAGATGTTAAATATATAGTAAGTATAATGAATAATAA
GTACAATGAAAAATACAATTATATTGAAAGGTTAAACTGGCAAGCATGGATGACATATTA
ATATTGTCAGAATTAAAGTGAACCAAAGTGTATCGAGAGGTTTCAAGTATCTTGAGTT
TCATGGCAAAGTGTAACTAGTTACTACAATGTTGGTGTGAAATTATCTGC
CTGGTGTGTCACACAAGTCTTACTGGAATAATTACTGGTAC

FIGURE 122

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58747
<subunit 1 of 1, 336 aa, 1 stop
<MW: 36865, pI: 9.15, NX(S/T): 2
MAVATAAAVLAALGGALWLAARRFVGPRVQRLRRGGDPGLMHGKTVLITGANSGLGRATAAE
LLRLGARVIMGCRDRARAEEAAGQLRRELQAAECGPEPGVSGVGELIVRELDLASLRSVRA
FCQEMLQEEPRLDVLIINNAGIFQCPYMKTEDGFEMQFGVNHLGHFLLTNLLLGLLKSSAPSR
IVVVSSKLYKYGDINFDDLNSEQSYNKSFCYSRSKLANILFTRELARRLEGTVNVTVLHPG
IVRTNLGRHIHIPLLVKPLFNLVSWAFFKTPVEGAQTSIYLASSPEVEGVSGRYFGDCKEEE
LLPKAMDESVARKLWDISEVMVGLLK
```

Important features:

Signal peptide:

amino acids 1-21

Short-chain alcohol dehydrogenase family protein

amino acids 134-144, 44-56 and 239-248

N-glycosylation site.

amino acids 212-215 and 239-242

FIGURE 123

GGGGATTGTAAAGAGGAAGNACTGTGCCAAAGNTATGGATGAATCTGTTGCAAGAAAATTN
TGGGATATCAGTGAAGTGATGGTTNGCCTGCTAAAATAGGAACAAGGAGTAAAAGAGCTGTT
TATAAAACTGCATATCAGTTATATCTGTGATCAGGAATGGTGTGGATTGAGAACTTGTACT
TGAAGAAAAAGAATTGGATATTGGAATAGCCTGNTAAGAGGNACATGTGGGTATTTGGAG
TTACTGAAAAATTATTTGGATAAGAGAATTCAGCAAAGATGTTAAATATATAGT
AAGTATAATGAATAATAAGTACAATGAAAATACAATTATATTGTAAAATTATAACTGGGCA
AGCATGGATGACATATTAATATTGTCAGAATTAAGTGACTCAAAGTGCTATCGAGAGGTTT
TTCAAGTATCTTGAGTTCATGGCAAAGTGTAACTAGTTACTACAATGTTGGTGT
TGTGTGGAAATTATCTGCCTGGCTT

FIGURE 124

GAGAGGACGAGGTGCCGCTGGAGAATCCTCCGCTGCCGCTGGCTCCGGAGCCAGCC
CTTCCTAACCAACCAACCTAGCCAGTCCAGCCAGCGCCTGTCACGGAC
CCCAGCGTTACC**ATG**CATCCTGCCGTCTCCTATCCTTACCGACCTCAGATGCTCCCTCT
GCTCCTGGTAACCTGGGTTTTACTCCTGTAACAACGTGAAATAACAAGTCTTGCTACAGAGA
ATATAGATGAAATTAAACAATGCTGATGTTGCTTAGTAAATTATGCTGACTGGTGT
CGTTTCAGTCAGATGTTGCATCCAATTTGAGGAAGCTCCGATGTCATTAAGGAAGAATT
TCCAAATGAAAATCAAGTAGTGTGCTGCCAGAGTTGATTGTGATCAGCACTCTGACATAGCCC
AGAGATAACAGGATAAGCAAATACCAACCCCTCAAATTGTTGTAATGGATGATGATGAAG
AGAGAATAACAGGGTCAGCGATCAGTGAAGCATTGGCAGATTACATCAGGCAACAAAAAG
TGACCCCATTCAAGAAATTGGACTTAGCAGAAATCACCCTTGTGATCGCAGCAAAAGAA
ATATCATTGGATATTGAGCAAAGGACTCGGACAACATAGAGTTGAACGAGTAGCG
AATATTTGCATGATGACTGTGCCTTCTTCTGCATTGGGATGTTCAAAACCGGAAAG
ATATAGTGGCGACAACATAATCTACAAACCACCGAGGCATTCTGCTCCGGATATGGTGTACT
TGGGAGCTATGACAAATTGATGACTTACAATTGGATTCAAGATAATGTGTTCTCTT
GTCCGAGAAATAACATTGAAAATGGAGAGGAATTGACAGAAGAAGGACTGCCTTTCTCAT
ACTCTTCACATGAAAGAAGATAACAGAAAGTTAGAAATATTCCAGAATGAAGTAGCTCGGC
AATTAATAAGTAAAAAGGTACAATAAACTTTTACATGCCATTGTGACAAATTAGACAT
CCTCTTCTGCACATACAGAAAACCTCCAGCAGATTGTCCTGTAATCGCTATTGACAGCTTAG
GCATATGTATGTGTTGGAGACTTCAAAGATGTATTAATTCTGGAAAACCTCAAGCAATTG
TATTGACTTACATTCTGGAAAACCTGCACAGAGAATTCCATCATGGACCTGACCCAACTGAT
ACAGCCCCAGGAGAGCAAGCCCAAGATGTAGCAAGCAGTCCACCTGAGAGCTCCTCCAGAA
ACTAGCACCCAGTGAATATAGGTATACTCTATTGAGGGATCGAGATGAGCTT**AAAAA**ACTTG
AAAAACAGTTGTAAGCCTTCAACAGCAGCATCACCTACGTGGTGAAATAGTAAACCTA
TATTTCATAATTCTATGTGTATTTTATTGAAATAACAGAAAGAAATTAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 125

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57689
<subunit 1 of 1, 406 aa, 1 stop
<MW: 46927, pI: 5.21, NX(S/T): 0
MHPAVFLSLPDLRCSLLLLVTWVFTPVTTEITSLATENIDEILNNADVALVNFYADWCRFSQ
MLHPIFEEASDVIKEEFPNENQVVFARVDCDQHSDIAQRYRISKYPTLKLFRNGMMMKREYR
GQRSVKALADYIRQQKSDPIQEIRDIAEITTLDRSKRNIIGYFEQKDSNDNYRVFERVANILH
DDCAFLSAFGDVSKPERYSGDNIIYKPPGHsapDMVYLGAMTNFDVTYNWIQDKCVPLVREI
TFENGEELTEEGLPFLILFHMKEDTESLEIFQNEVARQLISEKGTINFLHADCDKFRHPLLH
IQKTPADCPVIAIDSFRHMYVFGDFKDVLIPGKLKQFVFDLHSGKLHREFHHGPDPTDTAPG
EQAQDVASSPPESSFQKLAPSEYRYTLLDRDEL
```

Important features:

Signal peptide:

amino acids 1-29

Endoplasmic reticulum targeting sequence.

amino acids 403-406

Tyrosine kinase phosphorylation site.

amino acids 203-211

Thioredoxin family proteins

amino acids 50-66

FIGURE 126

ATTAAGGAAGAATTCCAAATGAAAATCAAGTAGTNTTGCCAGAGTNGATTGTGATCAGCA
CTCTGACATAGCCCAGAGATAACAGGATAAGCAAATACCCAACCCTCAAATTGTTCGTAATG
GGATGATGATGAAGAGAGAATACAGGGGTCAAGCGATCAGTGAAAGCATTGGCAGATTA

FIGURE 127

AGAGGCCCTCTGGAAAGTTGTCGGGCTGCCGCGCNGGAGCCGGTCGAGAGGACNAGG
TGCCGCTGCCTGGAGAACCTCCGCTGCCGTCGGCTCCGGAGCCCAGCCCTTCTAACCC
AACCCAAACCTAGCCCNGTCCCAGCCGCCAGCGCCTGTCCTGTCNCGGANCCAGCGTNACC
ATGCATCCTGCCGTCTCCTATCCTAACCGACCTCAGATGCTCCCTCTGCTCCTGGTAAC
TTGGGTTTTACTCCTGTAACAACGTAAATAACNNGTCTGATAACNNAGAATATAGATGAAA
TTTAAACNATGCTGATGTGGCTTAGTCATTATGCTGACTGGTGTGTTAGTCAG
ATGTGGCATCCAATTGGAGGANGCTCCGATGTCATTAAGGAAGAATTCCAAATGAAAA
TCAAGTAGTGTGTTGCCAGAGTTGATTGTGATCAGCACTCTGACATAGCCCAGAGATACAGGA
TAAGCAAATACCCAACCCCTCAAATTGTTGTAATGGGATGATGATGAAGAGAGAATACAGG
GGTCAGCGATCAGTGAAGCATTGGCAGATTACATCAGGC

卷之三

FIGURE 128

GCCCACGCGTCCGATGGCGTTCACGTTCGCGGCCTCTGCTACATGCTGGCGCTGCTGCTCA
CTGCCCGCCTCATCTTCTTGCCTTGGCACATTATAGCATTGATGAGCTGAAGACTGAT
TACAAGAACCTATAGACCAGTGTAAATACCCCTGAATCCCCTGTACTCCAGAGTACCTCAT
CCACGCTTCCTTGTGTATGTTCTTGAGCAGAGTGGCTTACACTGGGTCTCAATA
TGCCCCTCTGGCATATCATATTGGAGGTATATGAGTAGACCAGTGATGAGTGGCCCAGGA
CTCTATGACCCCTACAACCATCATGAATGCAGATATTCTAGCATATTGTCAAGGAAGGATG
GTGCAAATTAGCTTTATCTTCTAGCATTCTTACTACCTATATGGCATGATCTATGTT
TGGTGAGCTCTTAGAACAAACACACAGAAGAATTGGTCCAGTTAAGTGCATGCAAAAGCCAC
CAAATGAAGGGATTCTATCCAGCAAGATCCTGTCCAAGAGTAGCCTGTGGAATCTGATCAGT
TACCTTAAAAAATGACTCCTTATTTAAATGTTCCACATTTGCTGTGGAAAGACTG
TTTCATATGTTACTCAGATAAAGATTAAATGGTATTACGTATAAAATTAAATATAAAAT
GATTACCTCTGGTGTGACAGGTTGAACCTGCACCTTAAGAACAGCCATAATCCTCTG
AATGATGCATTAATTACTGACTGTCCAGTACATTGGAAGCTTTGTTATAGGAACCTGTA
GGGCTCATTTGGTTCATGGAAACAGTATCTAATTATAAATTAGCTGTAGATATCAGGTGC
TTCTGATGAAGTAAAATGTATATCTGACTAGTGGAAACTTCATGGGTTCCATCTGTC
ATGTCGATGATTATATGGATACATTACAAAAATAAAAGCGGGATTTCCTCGCTT
GAATATTATCCCTGTATATTGCATGAATGAGAGATTCCATATTCCATCAGAGTAATAAA
TATACTTGCTTAATTCTAACAGATAAGTAAACATGATATAAAATATGCTGAATTACTT
GTGAAGAACATTTAAAGCTATTAAATGTGTTTATTGTAAGACATTACTTATTAAG
AAATTGGTTATTATGCTTACTGTTCTAATCTGGTGGTAAAGGTATTCTTAAGAATTGCAGG
TACTACAGATTTCAAAACTGAATGAGAGAAAATTGTATAACCATCCTGCTGTTCTTAGT
GCAATACAATAAAACTCTGAAATTAAAGACTC

FIGURE 129

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA23330
<subunit 1 of 1, 144 aa, 1 stop
<MW: 16699, PI: 5.60, NX(S/T): 0
MAFTFAAFCYMLALLTAALIFFAIWHIIAFDELKTDYKNPIDQCNLTNPLVLPEYLIHAFF
CVMFLCAAELTLGLNMPLLAYHIWRYMSRPVMSGPGLYDPTTIMNADILAYCQKEGWCKLA
FYLLAFFYYLYGMIYVLVSS
```

Important features:

Signal peptide:

amino acids 1-20

Type II transmembrane domain:

amino acids 11-31

Other transmembrane domain:

amino acids 57-77 and 123-143

FIGURE 130

ATTATAGCATTGATGAGCTGAAGACTGATTACAAGATCCTATAGACCAGTGTAAATACCCCTG
AATCCCCTTGTACTCCCAGAGTACCTCATCCACGCTTCTTGTGTATGTTCTTGTGC
AGCAGAGTGGCTTACACTGGGTCTCAATATGCCCTCTGGCATATCATATTGGAGGTATA
TGAGTAGACCAGTGATGAGTGGCCAGGACTCTATGACCTACAACCACATGAATGCAGAT
ATTCTAGCATATTGTCAGAAGGAAGGATGGTCAAATTAGCTTTATCTTCTAGCATT
TTACTACCTATATGGCATGATCTATGTTGGTGAGCTCTAGAACACACAGAAGAATT
GGTCCAGTTAAGTGCATGCAAAAGCCACAAATGAAGGGATTCTATCCAGCAAGATCCTGT
CCAAGAGTAGCCTGTGGAATCTGATCAGTTACTTAAAAAATG

FIGURE 131

CGGACGCGTGGGGAAACCCCTCCGAGAAAACAGCAACAAGCTGAGCTGCTGTGACAGAGGG
GAACAAGATGGCGCGCCGAAGGGAGCCTCTGGGTGAGGACCCAACTGGGCTCCGCCGC
TGCTGCTGCTGACCATGGCCTTGGCCGGAGGTCGGGGACCGCTCGGCTGAAGCATTGAC
TCGGTCTGGGTGATA CGCGTCTGCCACCGGGCTGTCAGTTGACCTACCCCTTGCACAC
CTACCCCTAAGGAAGAGGAGTTGTACGCATGTCAGAGAGGTTGCAGGCTGTTCAATTGTC
AGTTTGTGGATGATGGAATTGACTTAAATCGAACTAAATTGGAATGTGAATCTGCATGTACA
GAAGCATATTCCAATCTGATGAGCAATATGCTGCCATCTGGTGCAGAACATCAGCTGCC
ATTGCGTGAAC TGAGACAAGAACAACTTATGTCCTGATGCCAAAATGCACCTACTCTTC
CTCTAACTCTGGT GAGGT CATTCTGGAGTGACATGATGGACTCCGACAGAGCTTCATAACC
TCTTCATGGACTTTTATCTTCAAGCCATGACGGAAAAATAGTTATATTCCAGTCTAACCC
AGAAAATCCAGTACGCACCACATTGGAGCAGGAGCCTACAAATTGAGAGAATCATCTCTAA
GCAAAATGTCCTATCTGCAAATGAGAAATTCAAGCGCACAGGAATTTCTTGAAGATGGA
GAAAGTGATGGCTTTAAGATGCCTCTCTTAAC TCTGGTGGATTTAACTACAAC TCT
TGT CCTCTCGGTGATGGTATTGCTTGGATTGTTGTGCAACTGTTGCTACAGCTGTGGAGC
AGTATGTTCCCTCTGAGAAGCTGAGTATCTATGGT GACTTGGAGTTATGAATGAACAAAAG
CTAAACAGATATCCAGCTTCTTCTTGTGGTTGTTAGATCTAAACTGAAGATCATGAAGA
AGCAGGGCCTCTACCTACAAAAGTGAAATCTGCTCATTCTGAAATTTAAGCATTTCTTT
AAAAGACAAGTGTAATAGACATCTAAAATTCCACTCCTCATAGAGCTTTAAAATGGTTCA
TTGGATATAGGCCTTAAGAAATCACTATAAAATGCAAATAAGTTACTCAAATCTGTG

FIGURE 132

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA26847
<subunit 1 of 1, 323 aa, 1 stop
<MW: 36223, pi: 5.06, NX(S/T): 1
MAAPKGSLWVRTQLGLPPLLLTMALAGGSGTASAEAFDSVLGDTASCHRACQLTYPLHTYP
KEEELYACQRGCRLFSICQFVDDGIDLNRKLECESACTEAYSQSDEQYACHLGCQNQLPFA
ELRQEQLMSLMPKMHLLFPLTLVRSFWSDMMDSAQSFITSSWTFYLQADDGKIVIFQSKPEI
QYAPHLEQEPTNLRESSLSKMSYLQMRNSQAHRFLEDGESDGFLRCLSLNSGWILTTLVL
SVMVLLWICCATVATAVEQYVPSEKLSIYGDLEFMNEQKLNRYPASSLVVVRSKTEDHEEAG
PLPTKVNLAHSEI
```

Important features:

Signal peptide:

amino acids 1-31

Transmembrane domain:

amino acids 241-260

N-glycosylation site.

amino acids 90-93

FIGURE 133

TTGGGTGATACGGCGTCTGCCACCGGGCCTGTCAGTTGACCTACCCCTGCACACCTACCC
TAAGGAAGAGGAGTTGTACGCATGTCAGAGAGGTTGCAGGCTGTTTCAATTGTCAGTTG
TGGATGATGGAATTGACTTAAATCGAACTAAATTGGAATGTGAATCTGCATGTACAGAAGCA
TATTCCCAATCTGATGAGCAATATGCTTGCCATCTGGTGCAGAATCAGCTGCCATTGCG
TGAACGTGAGACAAGAACAACTTATGTCCCTGATGCCAAAATGCACCTACTCTTCCTCTAA
CTCTGGTGAGGTGATTCTGGAGTGACATGATGGACTCCGC

FIGURE 134

CACACTGGCCGGATCTTTAGAGTCCTTGACCTTGACCAAGGGTCNGGAAAACAGCAACAA
GCTGAGCTGCTGTGACAGAGGGAAACAAGATGGCGCGCCGAAGGGAGCCTTGGGTGAGGAC
CCAAGTGGGCTCCCGCCGCTGCTGCTGACCATGGCCTGGCGAGGTTGGGGACCG
CTTCGGCTGAAGCATTGACTCGGTCTTGGGTGATACGGCGTCTGCCACCGGGCTGTCAG
TTGACCTACCCCTGCACACCTACCCCTAAGGAAGAGGAGTTGTACGCATGTCAGAGAGGTTG
CAGGCTGTTTCAATTGTCAGTTGTGGATGATGGAATTGACTTAAATCGAACTAAATTGG
AATGTGAATCTGCATGTACAGAACATATTCCAATCTGATGAGCAATATGCTGCCATCTT
GGTTGCCAGAACATCAGCTGCCATTGCTGAACTGAGACAAGAACAACTATGTCCTGATGCC
AAAAATGCACCTACTCTTCCTCTAACTCTGGTGAGGTATTCTGGAGTGACATGATGGACT
CCGC

FIGURE 135

GCAGGGTGGCAGTCGCTGAGAGGCAGGGAGGGCCGAGGCAGGGCTGGAGGCAGGGCCGGAGGT
GGGGCGCCGCTGGGCGCCGCACGGGCTTCATCTGAGGGCGCACGGCCCGCAGCCGAGC
GTGCGGACTGGCCTCCAAGCGTGGGCACAGAGCTGCCAGGAGCTGCAATGGGCCGGCTG
GGGATTCTTGTGCTGCCCTGGCGCCGTGTGGCTGCTCAGCTGGCCACGGAGAGGAGC
AGCCCCGGAGACAGCGCACAGAGGTGCTTCTGCCAGGTTAGTGGTTACTGGATGATTGT
ACCTGTGATGTTGAAACCATTGATAGATTTAAACTACAGGCTTCCAAAGACTACAAAA
ACTTCTTGAAAGTGACTACTTAGGTATTACAAGGTAACCTGAAGAGGCCGTGTCCTTCT
GGAATGACATCAGCCAGTGTGAAAGAAGGGACTGTGCTGTCAAACCATGTCAATCTGATGAA
GTTCCGTGATGAAATTAAATCTGCGAGCTACAAGTATTCTGAAGAAGCCAATAATCTCATTGA
AGAATGTGAACAAGCTGAACGACTTGGAGCAGTGGATGAATCTGTGAGTGAGGAAACACAGA
AGGCTGTTCTTCAGTGGACCAAGCATGATGATTCTCAGATAACTCTGTGAAGCTGATGAC
ATTCACTCCCCCTGAAGCTGAATATGTAGATTGCTTCTTAATCCTGAGCGCTACACTGGTTA
CAAGGGACCAGATGCTGGAAAATATGGAATGTCATCTACGAAGAAAATGTTAACGCCAC
AGACAATTAAAAGACCTTAAATCCTTGGCTCTGGCAAGGGACAAGTGAAGAGAACACT
TTTACAGTTGGCTAGAAGGTCTCTGTGTAGAAAAAGAGCATTCTACAGACTTATATCTGG
CCTACATGCAAGCATTAAATGTGCATTGAGTGCAAGATATCTTACAAGAGACCTGGTTAG
AAAAGAAATGGGACACAACATTACAGAATTCAACAGCGATTGATGGAATTGACTGAA
GGAGAAGGTCCAAGAAGGCTTAAGAACTTGTATTCTACTTAATAGAACTAAGGGCTT
ATCCAAAGTGTACCATTCTCGAGCGCCAGATTTCACACTCTTACTGGAAATAAAATTC
AGGATGAGGAAAACAAATGTTACTTCTGGAAATACCTCATGAAATCAAGTCATTCTTGTG
CATTGATGAGAATTCACTTGTGGGATAAAAAGAGCACACAAACTAAAGGAGGA
CTTCTGACTGCATTAGAAATATTCAAGAATTATGGATTGTGGTTGTTAAATGTC
GTCTGTGGGAAAGCTTCAGACTCAGGGTTGGCACTGCTCTGAAGATCTTATTCTGAG
AAATTGATAGCAAATATGCCAGAAAGTGGACCTAGTTATGAATTCCATCTAACAGACAAGA
AATAGTATCATTCAACGCATTGGAAGAATTCTACAAGTGTGAAAGAATTAGAAAATC
TCAGGAACCTGTTACAGAATTCATTAAAGAAAACAAGCTGATATGTGCCTGTTCTGGAC
AATGGAGGCGAAAGAGTGGAAATTCACTAACAGGCATAATAGCAATGACAGTCCTAACCAA
ACATTATATAAGTTGCTTTGTAAAGGAGAATTATATTGTTAAGTAAACACATT
AAAAATTGTGTTAAGTCTATGTATAACTACTGTGAGTAAAGTAATACTTAATAATGTG
GTACAAATTAAAGTTAATATTGAATAAAAGGAGGATTATCAAATTAAAAAA
AAAAAAAAAAAAAAAAAAAAAA

FIGURE 136

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53974
<subunit 1 of 1, 468 aa, 1 stop
<MW: 54393, pI: 5.63, NX(S/T): 2
MGRGWGFLFGLLGAVWLLSSGHGEEQPPETAAQRCFCQVSGYLDDCTCDVETIDRFNNYRLF
PRLQKLLESODYFRYYKVNLRPCPFWNDISQCGRRDCAVKPCQSDEVPDGIKSASYKYSEEA
NNLIEECEQAERLGAVDESLSEETQKAVLQWTKHDDSSDNFCEADDIQSPEAEYVDLLNPE
RYTGYKGPDAWKIWNVIYEENCFKPQTIKRPLNPLASGQGTSEENTFYSWLEGLCVEKRAFY
RLISGLHASINVHLSARYLLQETWLEKKWGNITEFQQRFDGILTEGEGRRLKNLYFLYLI
ELRALSKVLPFFERPDFQLFTGNKIQDEENKMLLIEILHEIKSFPLHFDENSFFAGDKKEAH
KLKEDFRLHFRNISRIMDCVGCFKCRWGKLQTQGLGTALKILFSEKLIANMPESGPSYEFH
LTRQEIVSLFNAFGRISTSVKELENFRNLLQNIH
```

Important features:

Signal peptide:

amino acids 1-23

N-glycosylation site.

amino acids 280-283 and 384-387

Amidation site.

amino acids 94-97

Glycosaminoglycan attachment site.

amino acids 20-23 and 223-226

Aminotransferases class-V pyridoxal-phosphate

amino acids 216-222

Interleukin-7 proteins

amino acids 338-343

FIGURE 137

GCTGGAAATATGGATGTCATCTACGAGAAACTGTTTAAGCCACAGACAATTAAAAGACCTT
TAAATCCTTGGCTTCTGGTCAAGGGACAAGTGAAGAGNACACTTTTACAGTTGGCTAGAA
GGTCTCTGTGTAGAAAAAAGAGCATTCTACAGACTTATATCTGGCCTACATGCAAGCATTAA
TGTGCATTTGAGTGCAAGATATCTTTACAAGAGACCTGGTTAGAAAAGAAATGGGGACACA
ACATTACAGAATTNAACAGCGATTGATGGAATTTGACTGAAGGAGAAGGTCCAAGAAGG
CTTAAGAACTTGTATTTCTCTACTTAATAGAACTAAGGGCTTATCCAAAGTGTACCATT
CTTNGAGCGCCCAGATTTCAACTNTTACTGGAAATAAAATTCAAGGATGAGGNAAACAAAAA
TGTTACTTTGGAAATACATGAAATCAAGTCATTCCTTGCATTTGATGAGAATTCA
TTTTTTGCTG

2024 RELEASE UNDER E.O. 14176

FIGURE 138

CGGACGCGTGGCGGACGCGTGGCGGACGCGTGGGTTGGAGGGGCAGGATGGGAGGGAA
AGTGAAGAAAACAGAAAAGGAGAGGGACAGAGGCCAGAGGACTTCTCATACTGGACAGAAC
CGATCAGGCATGGAACCTCCCTTCGTCACTCACCTGTTCTGCCCTGGTGTTCCTGACAGG
TCTCTGCTCCCCCTTAAACCTGGATGAACATCACCCACGCCTATTCCAGGGCCACCAGAAG
CTGAATTGGATACAGTGTCTTACAACATGTTGGGGTGGACAGCGATGGATGCTGGTGGC
GCCCTGGATGGCCTTCAGGCACCGGAGGGGACGTTATCGCTGCCCTGTAGGGGG
GGCCCACAATGCCCATGTGCCAAGGCCACTAGGTGACTACCAACTGGAAATTCATCTC
ATCCTGCTGTGAATATGCACCTGGGATGTCCTGTTAGAGACAGATGGTATGGGGGATT
ATGGTGAGCTAAGGAGAGGGTGGCAGTGTCTCTGAAGGTCCATAAAAGAAAAAGAGAA
GTGTGGTAAGGAAAATGGCTGTGTGGAGGGTCAAGGAGTTAAAACCTAGAAAGCAAA
AGGTAGGTAAATGTCAGGGAGTAGTCTTCATGCCCTTCACACTGGGAGCATGTTCTGAGGGT
GCCCTCCAAGCCTGGAGTAACTATTCATGCCCTTCACACTGGGAGCATGTTCTGAGGGT
CGTGCTTGTGGCAGCTCTGCTTCAGTTCTGGATATGTGCCGTGGATGCTTCATTCCA
GCCTCAGGGAAGCCTGGCACCCACTGCCAACGTGAGCCAGAGGAAGGCTGAGTACTGGTT
CCCAGAAGGAGATACTGGGTGGAAAAGATGGGCAAAGCGGTATGATGCCCTGGCAAAGGG
CCTGCATGGCTATCCTCATGCTACCTAATGTGCTTGCACACTGGTCTGAGGAGGCAAG
TTCAGACTCCTGCCAGGTGTGGCCACACCTGTAATTCTAGCAGTGGGAGGCAAG
GTGGGCAGATCAGTCAAGGAGTCAGGAGTTCAAGACCAGCCTGCCAACATGGTAAACTCCAT
CTCTACTAAAAAAAAAAACAAAAATTAGCTGGGTGCGCTAGTGCATGCCGTAACTCTC
ATCTACTCGGGAGGCTAACAGACAGGAGACTCTCACTCAACCCAGGAGGTGGAGGTTGCCGTG
AGCCAAGATTGTGCCCTGCACTCTAGCGTGGGTGACAGAGTAAGCGAGACTCCATCTCAA
AATAATAATAATAATTCAACTCCTTATCAGGAGTCCATGATCTGGCCTGGCACAGTAA
CTCATGCCGTAAATCCCAACATTGGGAGGCCAACGCAGGAGGATTGCTTGAGGTCTGGAG
GTTTGAGACCAGCCTGGCAACATAGAAAGACCCATCTCTAAATAATGTTAAAAAT

FIGURE 139

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57039

><subunit 1 of 1, 124 aa, 1 stop

><MW: 13352, pI: 5.99, NX(S/T): 1

MELPFVTHLFLPLVFLTGLCSPFNLDEHHPRLFPGPPEAEFGYSVLQHVGGGQRWMLVGAPW
DGPSGDRRGDVYRCPVGGAHNAPCAKGHLG DYQLGNSSHPAVNMHLGMSLLETGDGGFMVS

Important features:

Signal peptide:

amino acids 1-22

Cell attachment sequence.

amino acids 70-73

N-glycosylation site.

amino acids 98-101

Integrins alpha chain proteins

amino acids 67-81

FIGURE 140

CACAGTTCCCCACCATCACTCNTCCCATTCCCTCCAACTTATTTTAGCTTGCATTGGGA
GGGGGCAGGGATGGGAGGGAAAGTGAAGAAAACAGAAAAGGAGAGGGACAGAGGCCAGAGGAC
TTCTCATACTGGACAGAAACCGATCAGGCATGGAACCTCCCTCGTCACTCACCTGTTCTG
CCCCTGGTGTTCCTGACAGGTCTCTGCTCCCCCTTAACCTGGATGAACATCACCCACGCCT
ATTCCCAGGGCCACCAGAAGCTGAATTGGATACAGTGTCTTACAACATGTTGGGGTGGAC
AGCGATGGATGCTGGTGGCGCCCCCTGGATGGGCCTTCAGGCGACCGGAGGGGGACGTT
TATCGCTGCCCTGTAGGGGGGCCACAATGCCCATGTGCCAAGGGCCACTTAGGTGACTA
CCAACCTGGAAATTCACTCATCCTGCTGTGAATATGCACCTGGGATGTCTCTGTTAGAGA
CAGATGGTGATGG

FIGURE 141

AAAGTTACATTTCTCTGGAACCTCCTAGGCCACTCCCTGCTGATGCAACATCTGGGTTG
GGCAGAAAGGAGGGTGCTCGGAGCCGCCCTTCTGAGCTCCTGGGCCGGCTCTAGAACAA
ATTCAGGCTCGCTCGACTCAGACCTCAGCTCCAACATATGCATTCTGAAGAAAGATGGCT
GAGATGGACAGAACATGCTTATTTGGAAAGAAACAAATGTTCTAGGTCAAACACTGAGTCTACCA
AATGCAGACTTCACAATGGTTCTAGAAGAAATCTGGACAAGTCTTTCATGTGGTTTCT
ACGCATTGATTCCATGTTGCTCACAGATGAAGTGGCATTCTGCCTGCCCTCAGAACCTC
TCTGTACTCTCAACCAACATGAAGCATCTCTGATGTGGAGGCCAGTGATCGCGCTGGAGA
AACAGTGTACTATTCTGTCGAATACCAGGGGGAGTACGAGAGCCTGTACACGAGCCACATCT
GGATCCCAGCAGCTGGTGCTCACTCACTGAAGGTCTGAGTGTGATGTCACTGATGACATC
ACGGCCACTGTGCCATACAACCTCGTGTAGGGCACATTGGGCTCACAGACCTCAGCCTG
GAGCATCCTGAAGCATCCCTTAATAGAAACTCAACCATCCTTACCCGACCTGGATGGAGA
TCACCAAAGATGGCTTCCACCTGGTTATTGAGCTGGAGGACCTGGGGCCAGTTGAGTTC
CTTGTGGCTACTGGAGGGAGGCTGGTCCAGGAACATGTCAAAATGGTGAGGAGTGG
GGGTATTCCAGTGCACCTAGAAACCATGGAGCCAGGGCTGCATACTGTGTGAAGGCCAGA
CATTCGTGAAGGCCATTGGGAGGTACAGCGCCTTCAGCCAGACAGAACATGTGTGGAGGTGCAA
GGAGAGGCCATTCCCCTGGTACTGCCCTGTTGCCTTGTGGCTTATGCTGATCCTTGT
GGTCGTGCCACTGTTGCTGGAAAATGGGCCGGCTGCTCCAGTACTCCTGTTGCCCGTGG
TGGTCCTCCCAGACACCTTGAAAATAACCAATTCACCCAGAAGTTAATCAGCTGCAGAAGG
GAGGAGGTGGATGCCCTGTGCCAGGCTGTGATGTCTCTGAGGAACCTCCTAGGGCTGGAT
CTCATAGGGTTGGAGGGCCAGGTGAAGCCGAGAACCTGGTCTGCATGACATGGAAACC
ATGAGGGGACAAGTTGTGTTCTGTTTCCGCCACGGACAAGGGATGAGAGAACAGTAGGAAGA
GCCTGTTGTCTACAAGTCTAGAAGCAACCACAGAGGCTAGGGTGGTTGTCTAACAGAACAC
TGACTGAGGCTTAGGGGATGTGACCTCTAGACTGGGGCTGCCACTTGCTGGCTGAGCAACC
CTGGAAAAGTGACTTCATCCCTCGGTCTAACAGTTCTCATCTGTAATGGGGAATTACC
TACACACCTGCTAAACACACACACAGAGTCTCTCTATATACACACGTACACATAAA
TACACCCAGCACTGCAAGGCTAGAGGGAAACTGGTACACTCTACAGTCTGACTGATTAG
TGTTCTGGAGAGCAGGACATAATGTATGATGAGAACATGATCAAGGACTCTACACACTGGGT
GGCTTGGAGAGGCCACTTCCCAGAATAATCCTTGAGAGAAAAGGAATCATGGGAGCAATGG
TGTTGAGTTCACTCAAGGCCAATGCCGGTGCAGAGGGGAATGGCTTAGCGAGCTCTACAGT
AGGTGACCTGGAGGAAGGTACAGCCACACTGAAAATGGATGTGCATGAACACGGAGGATC
CATGAACACTGTAAAGTGGACAGTGTGACACTGCAGACAGCAGGTGAAATGTATGT
GTGCAATGCGACGAGAACATGCAGAACATGTCAGTAACATGTGCATGTTGTCCTTTTTC
TGTTGGTAAAGTACAGAAATTCAAGAAATAAAAGGCCACCTGGCAAAAGCGGTAAAAAA
AAAAAAAAAA

FIGURE 142

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57033
<subunit 1 of 1, 311 aa, 1 stop
<MW: 35076, PI: 5.04, NX(S/T): 2
MQTFTMVLEEIWTSFLMWFFYALIPCLLTDEVAILPAPQNLSQLSTNMKHLLMWSPVIAPGE
TVYYSVHEYQGEYESLYTSHIWIPSSWCSLTEGPECDVTDDITATVPYNLRVRATLGSQTSAW
SILKHPFNRNSTILTRPGMEITKDGFLVIELEDLGPQFEFLVAYWRREPGAEHVKMVRSG
GIPVHLETMEPGAAYCVKAQTFVKAIGRYSAFSQTECVEVQGEAIPLVLALFAFVGFMLILV
VVPLFVWKMGRLLQYS CCPVVLPDTLKITNSPQKLISCRREVDACATAVMSPEELLRAWIS
```

Important features:

Signal peptide:

amino acids 1-29

Transmembrane domain:

amino acids 230-255

N-glycosylation site.

amino acids 40-43 and 134-137

Tissue factor proteins.

amino acids 92-119

Integrins alpha chain proteins

amino acids 232-262

FIGURE 143

TCCTGCTGATGCACATCTGGTTGGCAAAAGGAGGTTGCTCGAGCCGCCCTTAGCTT
CCTGGCCGGCTCTAGAACAAATTCAAGGCTTCGCTGCGACTAGACCTCAGCTCCAACATATGCA
TTCTGAAGAAAGATGGCTGAGATGACAGAACATGCTTATTTGGAAAGAAACAATGTTCTAGG
TCAAACGTGAGTCTACCAAATGCAGACTTCACAATGGTTCTAGAAGAAATCTGGACAAGTCT
TTTCATGTGGTTTCTACGCATTGATTCCATGTTGCTCACAGATGAAGTGGCCATTCTGC
CTGCCCTCAGAACCTCTCTGTACTCTCAACCAACATGAAGCATCTCTGATGTGGAGCCCA
GTGATCGCGCTGGAGAAACAGTGTACTATTCTGTCGAATACCAGGGGAGTACGAGAGCCT
GTACACGAGCCACATCTGGATCCCCAGCAGCTGGTGCTCACTCACTGAAGGTCTGAGTGTG
ATGTCACTGATGACATCACGGCCACTGTGCCATACAAACCTTGTGTAGGGCACATTGGC
TCACAGACCTCAGCCTGGAGCATCCTGAAGCATCCCTTAATAGAAACTCAACCATCCTTAC
CCGACCTGGATGGAGATCACCAAAAGATGGCTNCACCTGGTTATTGAGCTGGAGGACCTGG
GGCCCCAGTTGAGTTCTGTGGCTANTGGAGGAGGGCGAACCCCTGCGGCGCAAGGG
GTTNGCGAACCCCTTGCAGGCCCTGGGTATCTCTCGAGAAAAGAGAGAGGCCAATATGACCC
ACATACTCAATATGGACGAANTGCTATTGTCCACCTGTTGAGTGGCGCTGGTTGAT

FIGURE 144

CCCACGCGTCCGCCACGCGTCCGAGGGACAAGAGAGAAGAGAGACTGAAACAGGGAGAAGA
GGCAGGAGAGGAGGGAGGTGGGGAGAGCACCAAGCTGGAGGCCGACACTGAGGGAGGGCGGG
GGAGGTGAAGAAGGGAGAGAGGGAGAAGAGGGCAGGAGCTGAAAGGGAGAGGGAGGAGGAG
GAGGAGATGCGGGATGGAGACCTGGAGTTAGGTGGCTGGAGAGCTTAATGAAAAGAGAAC
GGAGAGGAGGTGTGGGTTAGGAACCAAGAGGTAGCCCTGTGGCAGCAGAAGGCTGAGAGGA
GTAGGAAGATCAGGAGCTAGAGGGAGACTGGAGGGTTCCGGAAAAGAGCAGAGGAAAGAGG
AAAGACACAGAGAGACGGGAGAGAGAAGAAGAGTGGTTGAAGGGCGGATCTCAGTCCTG
GCTGCTTGGCATTGGGAACTGGGACTCCCTGTGGGAGGGAGAGGAAAGCTGGAAGTCCT
GGAGGGACAGGGTCCCAGAAGGGAGGGACAGAGGAGCTGAGAGAGGGGGCAGGGCGTGG
CAGGGTCCCTCGGAGGCCTCCTGGGATGGGGCTGCAGCTCGTCTGAGCGCCCTCGAGC
GCTGGTACTCTGGCTGCAGCTGGGGCAGCAGTCACATCGGACCAGCACCTGACCCCGAGG
ACTGGTGGAGCTACAAGGATAATCTCCAGGGAAACTCGTGCCAGGGCCTCCTTCTGGGC
CTGGTGAATGCAGCGTGGAGTCTGTGTGCTGTGGGAAGCGGCAGAGCCCCGTGGATGTGGA
GCTGAAGAGGGTTCTTATGACCCCTTCTGCCCTTAAGGCTCAGCACTGGAGGAGAGA
AGCTCCGGGAACCTTGTACAACACCGGCCACATGTCTCCTCCTGCCTGCACCCGACCT
GTGGTCAATGTGTCTGGAGGTCCCCCTCTTACAGCCACCGACTCAGTGAAC TGCGGCTGCT
GTTGGAGCTCGCGACGGAGCCGGCTCGAACATCAGATCAACCACCAAGGGCTCTGCTG
AGGTGCAGCTCATTCAACCAGGAACCTACGGGAATTCAAGCGCTGCCTCCCGCGC
CCCAATGGCCTGGCCATTCTCAGCCTCTTGTCAACGTTGCCAGTACCTCTAACCCATTCT
CAGTCGCCTCTTAACCGCGACACCACACTCGCATCTCCTACAAGAATGATGCCTACTTC
TTCAAGACCTGAGCCTGGAGCTCTGTTCCCTGAATCCTCGGCTCATCACCTATCAGGGC
TCTCTCAGCACCCGCCCTGCTCCGAGACTGTCACCTGGATCCTCATTGACCAGGGCCTCAA
TATCACCTCCCTCAGATGCACCTCCCTGAGACTCCTGAGCCAGAATCCTCCATCTCAGATCT
TCCAGAGCCTCAGCGGTAAACAGCCGGCCCTGCAGCCCTGGCCACAGGGCACTGAGGGC
AACAGGGACCCCCGGCACCCCGAGAGGGCGCTGCCAGGGCCCCAACTACCGCCTGCATGTGGA
TGGTGTCCCCCATGGTCGCTGA~~G~~ACTCCCCCTCGAGGATTGCACCCGCCGTCTAACCTC
CCCACAAGGCGAGGGAGTTACCCCTAAAACAAAGCTATTAAAGGGACAGAATACTTA

FIGURE 145

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA34353
<subunit 1 of 1, 328 aa, 1 stop
<MW: 36238, pI: 9.90, NX(S/T): 3
MGAAARLSAPRALVLWAALGAAAHHGPAPDPEDWWSYKDNLQGNFVPGPPFWGLVNAAWSLC
AVGKRQSPVDVELKRVLYDPFLPPLRLSTGGEKLRGTLYNTGRHVSFLPAPRPVVNVSGGPL
LYSHRLSELRLLFGARDGAGSEHQINHQGFSAEVQLIHFNQELYGNFSAASRGPNGLAILS
FVNVASTSNPFLSRLLNRTDTIRISYKNDAYFLQDLSLELLFPESFGFITYQGSLSTPPCSE
TVTWILIDRALNITSQMHSRLLSQNPPSQIFQSLSGNSRPLQPLAHRALRGNRDPRHPER
RCRGPNYRLHVDGVPHGR
```

Important features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 177-199

N-glycosylation site.

amino acids 118-121, 170-173 and 260-263

Eukaryotic-type carbonic anhydrases proteins

amino acids 222-270, 128-164 and 45-92

FIGURE 146

GGCGCCTGGTCTGCGCGTACTGGCTGTACGGAGCAGGAGCAAGAGGTCGCCGCCAGCCTCCGCCGCCAGCCTC
GTCGTGTCGGCCCTCGCTCTGCAGCTACTGCTCAGAAACGCTGGGGGCCACCCCTGGCAGACTAACGAA
GCAGCTCCCCCTCCCACCCCACTGCAGGTCTAATTGGACGCCATTCTCCAGGTTGAGGGAGC
CGCAGAGGCGGGAGGCTCGTGTATTCTGCAGTCAGCACCCACGCTGCCCGGACGCTCGGTGCTCAGGCCCTC
GCGAGCGGGCTCTCGCTGCCCTTGTGAAGGCTCTGGCGGCTGCAGAGGCCGTCGGTTGGCT
CACCTCTCCAGGAAACTTCACACTGGAGAGCCTAAAGGAGTGAAGAGCCTGTCTGGAGATTTCCTGGGAA
ATCCTGAGGTCAATTATATGAAGTGTACCGCGGGAGTGGCTCAGAGTAACCACAGTGTCTCATGGCTAGA
GCAATTCCAGCCATGGGTTCCAATGCCACTTATTGGAGAAAACCTTGGAAAAAATACATGGATGAGGATGGT
GAGTGGTGGATAGCCAACACGAGGGAAAGGGCATCACAGACAATGACATGCAGAGTATTGGACCTTCAT
AATAAATTACGAAGTCAGGTGTATCCAACAGCCTCTAATATGGAGTATATGACATGGATGTAGAGCTGGAAAGA
TCGAGAACCTGGCTGAAAGTTGCTGGAACATGGACCTGCAAGCTGCTTCATCAATTGGACAGAAAT
TTGGGAGCACACTGGGAAAGATATAGGCCCCCGACGTTCATGTACAATCGTGGTATGATGAAGTGAAGACTTT
AGCTACCCATATGAACATGAATGCAACCCATATTGTCCTTCAGGTGTTCTGGCCCTGTATGTACACATTATACA
CAGGTGCTGTGGCAACTAGTAACAGAACGTTGTGCCATTAAATTGTCATAACATGAACATCTGGGGCAG
ATATGGCCCAAAGCTGTCTACCTGGTGTGCAATTACTCCCCAAAGGAAACTGGTGGGCCATGCCCTAACAA
CATGGCGGCCCTGTTCTGCCACCTAGTTGGAGGGGCTGTAGAGAAAATCTGTGCTACAAAGAAGGG
TCAGACAGGTATTATCCCCCTCGAGAACAGGAAACAAATGAAATAGAACGACAGCAGTCACAAGTCCATGACACC
CATGTCGGACAAGATCAGATGATAGTAGCAGAAATGAAGTCATAAGCGCACAGCAAATGTCCCAAATTGTTCT
TGTGAAGTAAGATTAAGAGATCAGTGCAGAACACCTGCAATAGGTACGAATGTCCTGCTGGCTTTGGAT
AGTAAAGCTAAAGTTATTGGCAGTGTACATTATGAAATGCAATTCCAGCATTGCTAGAGCTGCAATTCAATTGTT
ATAATAGACAATGATGGCTGGTAGATATCACTAGAACAGGAAGAAAAGCATTATTCATCAAGTCCAAATAGA
AATGGTATTCAAACAAATTGGCAAATATCAGTCTGCTAATTCTCAGTCCTCTAAAGTAACAGTCAGGCTGTG
ACTTGTGAAACAACACTGTGAAACAGCTGTCCTTCATAAGCCTGCTTCACATTGCCAAAGAGTATACTGTCCT
CGTAACTGTATGCAAGCAAACATTGCTCGTAAATTGAAACTCGAGTTATTCTGATCTGTCAGTATC
TGAGCAGCAGTACATGCTGGAGTGGTCTGAAATCACGGTGGTTATGTTGATGTAATGCCCTGGAACAAAAGA
AAGACCTACATTGCTTTTCAAGATGAAATCTCTCAGAAAGTTACAGAACATTCCAGGAGGAAAGGCATT
AGAGTGGTGTGTTGTTGTTGTAACTGAATACTGGAGAGGACATAAAGACTATTCAAATGCAATATTCTGA
ATTGGTATAAAACTGTAACATTACTGTACAGACTACACTAATTTCAGCCAAAAGGTGCCAAATGCATA
TAAATCTGATAAAACAAAGTCTATAAAATAAAACATGGACATTAGCTTGGAAAAGTAATGAAAATATAATGG
TTTAGAAATCCTGTTAAATATTGCTATATTCTTAGCAGTTATTCTACAGTTAATTACATAGTCATGATT
GTTCTACGTTTCATATATTATGGTGTGTTGATATGCCACTAATAAAATGAATCTAAACATTGAATGTAATG
GCCCTCAGAAAATCATCTAGTGCATTAAAATAATCGACTCTAAACTGAAAGAACCTTATCACATTCCCC
AGTTCAATGCTATGCCATTACCAACTCCAAATACTCAAATAATTCCACTTAATAACTGTAAGTTTTTC
TGTTAATTAGGCATATAGAATATTAAATTCTGATATTGCACTCTTATTATATAAAATACTTAAATATC
CAAATGAATCTGTTAAATGTTGATTCCCTGGGAATGGCCTTAAAATAATGTAATAAAAGTCAGAGTGGTGGT
ATGAAAACATTCTAGTGTACATGTAGTAATGTAGGGTTAACATGGACAGGCCAGAGCTTCTATGTACTGTTA
AAATTGAGGTACATATTCTTTGTATCCTGGCAAATACTCTGCAAGGCCAGGAAGTATAATGCAAAAGTT
GAACAAAGATGAACATGTTACATTACCAATTGCCACTGATTTTTAAATGGTAAATGACCTTGATATAAA
ATATTGCCATATCATGGTACCTATAATGGTGTATATTGTTCTATGAAAATGTAATTGTGCTTGATACTAAA
AATCTGTAAGATGTTAGTTGGTAATTCTGCTGGTGGATTACATATTAAATTCTGCTGGTGGAA
TAAACATTAAAATTAATCATGTTCAAAAAAAAAAA

FIGURE 147

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45417
<subunit 1 of 1, 500 aa, 1 stop
<MW: 56888, pi: 8.53, NX(S/T): 2
MKCTAREWLRVTTVLFMARAIPAMVVPNATLLEKLYMDEGEWWIAKQRGKRAITDNDM
QSILDLNKLRSQVYPTASNMEYMTWDVELERSAESWAESCLWEHGPASLLPSIGQNLGAHW
GRYRPPTFHVQSWYDEVKDFSYPPYEHECNPYCPFRCSPVCTHYTQVVWATSNRIGCAINLC
HNMMNIWGQIWPKAVYLVNCNYSPKGNWWGHAPYKHGRPCSACPPSFGGGCRENLCYKEGSDRY
YPPREEETNEIERQQSQVHDTHVRTRSDSSRNEVISAQQMSQIVSCEVRLRDQCKGTTCNR
YECPAGCLDSKAKVIGSVHYEMQSSICRAAIHYGIIDNDGGWVDITRQGRKHYFIKSNRNGI
QTIGKYQSANSFTVSKVTVQAVTCETTVEQLCPFHKPASHCPRVYCPRNCMQANPHYARVIG
TRVYSDLSSICRAAVHAGVVRNHGGYVDVMPVDKRKYIASFQNGIFSESLQNPPGGKAFRV
FAVV
```

Important features:

Signal peptide:

amino acids 1-20

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7 protein

amino acids 165-186, 196-218, 134-146, 96-108 and 58-77

N-glycosylation site

amino acids 28-31

FIGURE 148

GC GGAGACAAGCGCAGAGCGCAGCGCACGCCACAGACAGCCCTGGCATCCACCGACGGCG
CAGCCGGAGCCAGCAGAGCCGGAGGCCGGCCGGCAGAGAAAGCCGAGCAGAGCTGGGT
GGCGTCTCCGGGCCGCGCTCCGACGGCCAGGCCCTCCCCATGTCCTGCTCCCACGCCG
CGCCCTCCGGTCAGCATGAGGCTCTGGCGGCCGCTGCTCCTGCTGCTGGCGCTGT
ACACCGCGGTGTGGACGGGTCAAATGCAAGTGCTCCGGAAGGGACCCAAGATCCGCTAC
AGCGACGTGAAGAAGCTGGAAATGAAGCCAAAGTACCCGACTGCGAGGAGAAGATGGTTAT
CATCACCAAGAGCGTGTCCAGGTACCGAGGTCAAGGAGCACTGCCTGCACCCCAAGCTGC
AGAGCACCAAGCGCTTCATCAAGTGGTACAACGCCCTGGAACGAGAAGCGCAGGGTCTACGAA
GAATAGGGTAAAAACCTCAGAAGGGAAAACTCCAACCAGTTGGGAGACTTGTGCAAAGGA
CTTGAGATTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCCTTTC
TTTCTCACAGGCATAAGACACAAATTATATATTGTTATGAAGCACTTTACCAACGGTCAG
TTTTACATTTATAGCTGCGTGCAGGCTTCCAGATGGGAGACCCATCTCTTGTGCT
CCAGACTTCATCACAGGCTGCTTTATCAAAAAGGGAAAACTCATGCCTTCCTTTAA
AAAATGCTTTTGTTAGGCTGCTGGGACACATTTCATTGAGTGTGCTCCATTCTAGCTGGGAAGC
TTCCGCTTAGAGGTCTGGCCTGGCACAGCTGCCACGGCTCTCTGGCTTATGGCG
GTCACAGCCTCAGTGTGACTCCACAGTGGCCCTGTAGCCGGCAAGCAGGAGCAGGTCTCT
CTGCATCTGTTCTTGAGGAACCTCAAGTTGGTTGCCAGAAAATGTGCTTCATTCCCCCT
GGTTAATTTTACACACCCTAGGAAACATTCCAAGATCCTGTGATGGCGAGACAAATGATC
CTTAAAGAAGGTGTGGGTCTTCCAACCTGAGGATTCTGAAAGGTTCACAGGTTCAATA
TTAATGCTTCAGAAGCATGTGAGGTTCCAACACTGTCAGCAAAACCTTAGGAGAAA
TAAAAATATGAATAACATGCGCAATAACACAGCTACAGACACACATTCTGTTGACAAGGGAA
AACCTCAAAGCATGTTCTTCCCTCACCAACAGAACATGCACTAAAGCAATATAT
TTGTGATTCCCCATGTAATTCTCAATGTTAACAGTGCAGTCCTCTTGAAAGCTAAGAT
GACCATGCGCCCTTCCTGTACATATAACCTTAAGAACGCCCTCCACACACTGCC
CAGTATATGCCGCATTGTAAGTGTGTTATGCTATGTACATGTCAGAAACCATTAGCAT
TGCATGCAGGTTCATATTCTTAAGATGGAAAGTAATAAAATATTTGAAATGTA
AAAAAA

FIGURE 149

MSLLP RRRAPPVSMRLLA AALLLLL ALY TARVDGSKCKCSRKGPKIRYSDVKKLEMKPKYPH
CEE KMVIITT KSVSRYRGQEHCLHPKLQSTKRFIKWYN A WNEKRRVYEE

Signal sequence:

amino acids 1-34

FIGURE 150

CCCCCAGGGACTGCTATGGCTTCTTGTGTTCACCCGGCTGCGTCATGTTAAACTCCAATGTCCTCTGTG
GTTAACTGCTCTGCCATCAAGTTCACCTCATTGACAGCCAAGCACAGTATCCAGTTGCAACACAAATTATGG
CAAATCCGGGCCTAAGAACACCGTTACCAATGAGATCTTGGGTCCAGTGGAGCAGTACTTACGGGCTCCCTA
TGCCTCACCCCCCACTGGAGAGAGGGCGGTTCAGCCCCAGAACCCCCGTCCTGGACTGGCATCCGAAATAC
TACTCAGTTGCTGCTGTGCCCCAGCACCTGGATGAGAGATCCTTACTGCATGACATGCTGCCCATCTGGTT
TACCGCCAATTGGATACTTGTGATGACCTATGTTCAAGATCAAATGAAGACTGCCTTACTTAAACATCTACGT
GCCAACGGAAGATGGAGCCAACACAAAGAAAACGCAGATGATAAACGAGTAATGACCGTGGTAAGACGAAGA
TATTGATGACAGAACAGTAAGAACGCCCCTGATGGTCTATTCATGGGGATCTTACATGGAGGGCACCGCAA
CATGATTGACGGCAGCATTGGCAAGCTACGGAAACGTCATCGTGATCACCATTAACTACCGTCTGGAAACT
AGGGTTTTAAGTACCGGTGACCAGGCAGCAAAGGAACATATGGCTCTGGATCAGATTCAAGCACTGCCGTG
GATTGAGGAGAATGTGGAGCCTTGGCGGGACCCCAAGAGAGTGAACATCTTGGCTCGGGCTGGGCCTC
CTGTGTCAGCCTGTTGACCTGTCCTACTACAGAAGGTCTTCCAGAAGGCCATCATTGAGCGGCACCGC
CCTGTCAGCTGGCAGTGAACCTACAGCCGCAAGTACACTCGGATATTGGCAGACAAGGTGGCTGCAACAT
GCTGGACACCACGGACATGGTAGAATGCCTGCGAACAGAAACTACAAGGAGCTCATCCAGCAGACCATCACCC
GGCACCTACCACATAGCCTCGGGCGGTGATCGACGGCAGTCATCCAGCAGACGACCCCCAGATCCTGATGGA
GCAAGGCAGATTCTCAACTACGACATCATGCTGGCGTCAACCAAGGGGAAGGCCTGAAGTCTGTGGACGGCAT
CGTGGATAACGAGGACGGTGTGACGCCAACGACTTTGACTTCTCGTGTCCAATTGTTGACAACTTACGG
CTACCTGAAGGGAAAGACACTTTGCGGGAGACTATCAAGTTGATGTACACAGACTGGCCGATAAGGAAACCC
GGAGACGCGCGAAAACCTGGTGGCTCTTTACTGACCACAGTGGGTGGCCCCGCGTGGCGCCGACCT
GCACGCGCAGTACGGCTCCCCCACCTACTTCTATGCTTCTATCATCACTGCCAAGCGAAATGAAGCCCAGCTG
GGCAGATTGGCCCATGGTGTGAGGTCCCCATGTTCTCGGATCCCCATGATCGGTCACCGAGCTTTCAG
TTGTAACCTTCCAAGAACGACGTGATGCTCAGCGCGTGGTGTGACCTACTGGACGAACCTCGCCAAAACCTGG
TGATCCAATCAACCAGTCTCAGGATACCAAGTTGATCATGACACAAACCCAAACCGCTTGAAGAAGTGGCTG
GTCAGTAAATCCAAAAGACAGCTCTGATATTGGCTGAAACCCAGAGTGGAGAGATCACTACCGGG
AACGAAAGTGGCTTGTGAAACTCGTCTCATTGACACATTTCCCTATGGCACCCGCGATCTCCGCAAGATATGGCC
AACACAAAGGTTCTCCACAGACATCCTGCCAACATCCAAAACACTCTAAGGACCTCACAAAACAGGGCTGA
GGACACAACGTCTCATTGAAACAAACGAGATTATTCCACCGAATTAAAGTGTGACCTATTGCCGTGGCGTC
GCTCTCTTCTCAACATCTAGCTTGTGCGCTGTACTACAAAAAGGACAAGAGCGCATGAGACTCACAG
GCGCCCACTGGCAGAGAACACCACAAATGATATGCTCACATCCAGAACGAAGAGATCATGCTCTGAGAT
GAAGCAGCTGGAACACGATCACGAGTGTGAGTCGCTGCAGGCACACACTGAGGCTCACCTGCCGCCAGA
CTACACCCTACGCTGCCGGTGCAGATGACATCCACTTATGACGCCAACACCATCACCATTGATTCCAAA
CACACTGACGGGATGCGCTTGCACACTTTAACACCTTCAGTGGAGGACAAACAGTACAAATTACCCCA
CGGACATTCAACCACTAGAGTATAGCTTGCCTATTCCCTTCTATCCCTACCCGCTCAGCAACAT
AGAAGAGGAAAGGAAAGAGAGAGAGAGAGAGAGAAAGTCTCAGACCAGGAATGTTTGTCCACT
GACTTAAGACAAAATGCAAAAGCAGTCATCCCATCCGGCAGACCCCTATGTTGGTGTCCAGTATTAC
AAGATCAACTCTGACCTGTGAAATGTGAGAAGTACACATTCTGTTAAATAACTGTTAAGATCTTACCA
CTCCAATCAATGTTAGTGTGATAGGACATCACCATTCAAGGCCCCGGGTGTTCCACGTGATGGAAGCAGCT
GACACTTGTGAAACTCAGCCAAGGACACTTGATATTTTAATTACAATGAAAGTTAAACATTCTTCTGTG
CACACAATGGATGGCTCTCTTAAGTGAAGAAAGAGTCATGAGATTGGCCAGCACATGGAGCTGTAATCCAG
AGAGAAGGAAACGTAGAAATTATTAAAGAATGGACTGTGCAAGCGAAATCTGACGGTTCTGTGCAAAGAG
GTGTTTGCCAGCCTGAACTATATTAAAGAGACTTTGT

FIGURE 151

MLNSNVLLWLTALAIKFTLIDSQAQYPVVNTNYKIRGLRPLPNEILGPVEQYLGVPYASP
PTGERRFQPPEPPSSWTGIRNTTQFAAVCPQHLDERSLLHDMLPIWFTANLDTLMTYVQDQN
EDCLYLNIVPTEGANTKKNADDITSNDRGEDEDIHDAQNSKKPVMVYIHGGSYMEGTGNMI
DGSILASYGNVIVITINYRLGILGFLSTGDQAAKGNYGLLDQIQALRWIEENVGAFGGDPKR
VTIFGSGAGASCVSLLTLSHYSEGLFQKAIIQSGTALSSWAVNYQPAKYTRILADKVGCNML
DTTDMVECLRNKNYKELIQQTITPATYHIAFGPVIDGDVIPDDPQILMEQGEFLNYDIMLGV
NQGEGLKFVDGIVDNEDGVTNDFDFSVSNFVDNLYGYPEGKDTLRETIKFMYTDWADKENP
ETRRKTLVALFTDHQWVAPAVAADLHAQYGSPTYFYAFYHHCQSEMKPSWADSAHGDEVPYV
FGIPMIGPTELFSCNFSKNDVMLS A VVMTYWTNFAKTGDPNQPVPQDTKFIHTKPNRFE
WSKYNPKDQLYLHIGLKPRVRDHYRATKVAFWLELVPHLHN LNEIFQYVSTTKVPPDMTS
FPYGTRRSPAKIWPTTKRPAITPANNPKHSKDPHKTGPEDTVLIE
ASLLFLNILAFAALYYKKDKRRHETHRRPSQQRNTTNDIAHIQNEEIMSLQMKQLEHDHECE
SLQAHDTLRLTCPDYTLRRSPDDIPLMTPNTITMIPNLTGMQPLHTFNTFSGGQNSTN
LPHGHSTTRV

Signal sequence:

amino acids 1-24

Transmembrane domains:

amino acids 189-204, 675-692

FIGURE 152

GGGAAAGATGGCGGCGACTCTGGGACCCCTGGGTGTCGGCAGCAGTGGCGGCATGTTGT
CGGCTCGGGATGGGTCCAGGATGTTACTCCTTCTTGTGGGGCTGGCAGGGCCA
CAGCAAGTCGGGGGGTCAAACGTTCGAGTACTGAAACGGGAGCACTCGCTGTCGAAGCC
CTACCAGGGTGTGGCACAGGCAGTTCTCACTGTGGAATCTGATGGCAATGCCATGGTGA
TGACCCAGTATATCCGCCTACCCCAGATATGCAAAGTAAACAGGGTGCCTGTGGAACCGG
GTGCCATGTTCTGAGAGACTGGAGTTGCAAGGTCACTCAAAATCCATGGACAAGGAAA
GAAGAATCTGCATGGGATGGCTTGGCAATCTGGTACACAAAGGATGGATGCAGCCAGGGC
CTGTGTTGGAAACATGGACAAATTGTTGTGGGGCTGGGAGTATTGTAGACACCTACCCCAAT
GAGGAGAAGCAGCAAGAGCGGGTATTCCCTACATCTCAGGCATGGTGAACAACGGCTCCCT
CAGCTATGATCATGAGCGGGATGGCGGCCTACAGAGCTGGGAGGCTGCACAGCCATTGTCC
GCAATCTTCATTACGACACCTCCTGGTATTGCTACGTCAAGAGGCATTGACGATAATG
ATGGATATTGATGGCAAGCATGAGTGGAGGGACTGCATTGAAGTGCCGGAGTCCGCCTGCC
CCGCGGCTACTACTTCGGCACCTCCTCCATCACTGGGATCTCTCAGATAATCATGATGTCA
TTTCCTTGAAGTTGTTGAACTGACAGTGGAGAGAACCCAGAAGAGGAAAAGCTCCATCGA
GATGTGTTCTGCCCTCAGTGGACAATATGAAGCTGCCTGAGATGACAGCTCCACTGCCGCC
CCTGAGTGGCCTGCCCTTCCATCGTCTTTCTCCCTGGTGTGTTCTGTATTTGCA
TAGTCATTGGTATCATACTACAACAAATGGCAGGAACAGAGCCAAAGCGCTTCTACTTGA
GCCCTCCTGCTGCCACCACCTTGACTGTCACCCATGAGGTATGGAAGGAGCAGGCAGTG
GCCTGAGCATGCAGCCTGGAGAGTGTCTTGCTCTAGCAGCTGGTGGGGACTATATTCTG
TCACTGGAGTTTGAATGCAGGGACCCCGATTCCCATGGTGTGCATGGGACATCTAACT
CTGGTCTGGGAAGCCACCCACCCAGGGCAATGCTGCTGTGATGTGCCCTTCCCTGCAGTCC
TTCCATGTGGAGCAGAGGTGTGAAGAGAATTACGTGGTTGTGATGCCAAACACAGAAC
AGAATTTCATAGCCCAGGCTGCCGTGTTGACTCAGAAGGCCCTCTACTCAGTTTG
AATCCACAAAGAATTAAAACGGTAACACACCAGGGTTCTGACCATCCATTGTTGGGTT
TTGCATTGACCCAACCCTCTGCCTACCTGAGGAGCTTCTTGGAAACCAGGATGGAAACT
TCTTCCCTGCCTTACCTTCTTCACTCCATTGTCCTCTGTGTCACCTGAGCTG
GGAAAGGCATTGGATGCCCTCTGTTGGGCTGGGGCTGAGAACACACCTGCGTTCA
TGGCCTTCATTAGGTGCCCTAGGGAGATGGCTTCTGCTTGGATCACTGTTCCCTAGCAT
GGGTCTGGGTCTATTGGCATGTCCATGCCCTCCCAATCAAGTCTCTCAGGCCCTCAGTG
AAGTTGGCTAAAGGTTGGTAAAAATCAAGAGAACGCTGGAAGACATCATGGATGCCATG
GATTAGCTGTGCAACTGACCAAGCTCCAGGTTGATCAAACCAAAAGCAACATTGTATGTG
GTCTGACCATGTGGAGATGTTCTGGACTTGCTAGAGCCTGCTAGCTGCATGTTGTAGT
TACGATTTTGAATCCCACTTGAGTGTGAAAGTGTAAAGGAAGCTTCTTACACCTT
GGCTTGGATATTGCCAGAGAACAGGAAATTGGCTTTCTTAATGGACAAGAGACAGT
TGCTGTTCTCATGTTCCAAGTCTGAGAGCAACAGACCCCTCATCTGTGCCCTGGAAAGAGTT
CACTGTCATTGAGCAGCACAGCCTGAGTGCTGCCCTGTCAACCCTTATTCCACTGCCCTA
TTGACAAGGGGTACATGCTGTCACCTTACTGCCCTGGGATTAAATCAGTTACAGGCCAG
AGTCTCCTGGAGGGCCTGGAACACTGAGTCTCTATGAACCTCTGTAGCCTAAATGAAAT
TCTTAAATCACCGATGGAACCAAAAAAAAAAAAAAGGGCGGCCCGACTCTAGAGTCG
ACCTGCAGTAGGGATAACAGGGTAATAAGCTTGGCCGCATGG

FIGURE 153

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50911
><subunit 1 of 1, 348 aa, 1 stop
><MW: 39711, pI: 8.70, NX(S/T): 1
MAATLGPLGSWQQWRRCLSARDGSRMLLLLLLGSGQGPQQVGAGQTFEYLKREHSLSKPYQ
GVGTGSSSLWNLMGNAMVMTQYIRLTPDMQSKQGALWNRVPCFLRDWELQVHFKIHGQGKKN
LHGDGLAIWYTKDRMQPQPGPVFGNMDKFVGLGVFVDTYPNEEKQQERVFPYISAMVNNGSLSY
DHERDGRPTELGGCTAIVRNLHYDTFLVIRYVKRHTIMMDIDGKHEWRDCIEVPGVRLPRG
YYFGTSSITGDLSDNHDVISLKLFEVTERTPEEEKLHRDVFLPSVDNMKLPMTAPLPLS
GLALFLIVFFSLVFSVFAIVIGIILYNKWQEQRKRFY
```

Signal sequence:

amino acids 1-38

Transmembrane domain:

amino acids 310-329

FIGURE 154

CCGAGCCGGCGCGCAGCGACGGAGCTGGGGCCGGCTGGGACCATGGCGTGAGTGAATCTACGGATCAGTCT
CTGATGGTGGTCGTTAACCTCAGTGGGACTCCAAGATTCCATGAAGAAAATCAGTTGCTTCATTCAAGAAT
TGGGGTCTGGCTCAGAATTCTCGCAGCTGGTAAAAATCTGTTTCTAGAAAGAGGTTAATTAAATGCCTGCAGTCT
GACATGTTCCCGATTGAGGTGAAACCATGAAGAGAAAATAGAATACTTAATAATGCTTTCCGCAACCGCTTCT
TGCTGCTGCTGGCCCTGGCTGCCTGCTGCCATTGTGAGCCTCAGCCTGCAGTTCTCCACCTGATCCGGTGT
CGACTCCTAAGAATGGAATGAGTAGCAAGAGTCGAAAGAGAATCATGCCGACCCCTGTGACGGAGCCCCCTGTGA
CAGACCCGTTATGAAGCTCTTGTACTGCAACATCCCCAGTGTGGCCAGCGCAGCATGGAAGGTATGCC
CGCATCATTAAAGCTGGTCTCAGTCATGTGTCATTGCCACGGAGACAGGTACCCACTGTATGTCATCCCA
AAACAAAGCACCAGAAATTGACTGCACTCTGGGGTAACAGAAACCGTATCACCAAAATGGAAGCTTCA
TTAGTCACATGTCAAAAGGATCCGGAGCCTTCTGAAAGCCCTTGAACTCCTTGCCCTTTACCCAAATCACC
CATTGTGTGAGATGGGAGAGCTCACACAGACAGGAGTTGTGCAAGCATTGCAAGACGGTCAGCTGCTGAGGGATA
TCTATCTAAAGAAACACAAACTCCTGCCAATGATTGGTCTGCAGACCAGCTCTATTAGAGACCACTGGAAAAA
GCCGGACCCATAAAAGTGGCTGGCCTTGCCTTATGGCTTCTCCAGATTGACTGGAAGAAGATTATTCA
GGCACCAGCCAAGTGGCTGTTGCTCTGGAAGCTGCTATTGCCGGTAAGAAACCACTGATCTGAAAAGGAGC
AGCGTCGTCACTACCTCTACGTTGAAAAACAGCCAGCTGGAGAAGACCTACGGGAGATGCCAAGATCGTGG
ATGTCCTCACCAAGCAGCTAGAGCTGCCAACCCATAGACTCCATGCTCTGCCACTCTGCCACAATGTCAGCT
TTCCCTGTACCAGAAATGGCTGTGTTGACATGGAGCACTTCAGGTAATTAAAGACCCATCAGATGAGGATGAAA
GGGAAAGACGGGAGAAGAAATTGACTTCGGGTATTCTCTCTGGGTGCCACCCATCCTGAACCAAACCATCG
GCCGGATGCGTGCACCGAGGGCAGGAAAGAAGAGCTTTGCCCTCTACTCTGCTCATGATGTCACTCTGT
CACCAGTTCTCAGTCCTGGCCTTCAGAAGCCAGGTTCCAAGGTTGCAGCCAGGTTGATCTTGAGCTTT
GCCAAGACAGAGAAAAGCCCAGTGAACATTCCGTCGGATTCTTACAATGGCGTCATGTCACATTCCACACCT
CTTCTGCCAAGACCACCAAGCGTTCTCCAAGCCCATGTGCCGTTGAAAACTGGTCCGTTGTGAAAAA
GGGACATGTTGTAGGCTGGGTGGCAGTGTGACAAATTATTATGATGTCATGTCACAGGAAAGGATTCTAAAGG
TATGCACTACAGCAGTATAGAATCCATGCCAATACAGAGCATAGGAAAGGTCACCTCTAGTTGTCGTTAC
TAAGGGTAGAAGATTATTGCTTTAAAGGCTAAATATTGTTGTTGGAAACCAGATGGTGGGTTGAACAGT
AAGCACATTGCTGCAATGTGGTACGTGAATTGCTGGTACAAAATGGCCAGTTCACAGAGGAATAGAAGGTACTT
TATCATAGCCAGACTTCGCTTAGAATGCCAGAATAATATAGTCAAGACCTGAAGTTGCCAATCCAAGTTGAC
TCTTCTGGCCTGCCCATGTTACTATGTGATGGAACCGCACACCTCAACAAAATTTTAACTTCTAGACATT
TTTACCTTGTCTTGTAAAGAATTCTTGAAGTGATTATCTAAATAAGGTTGGCAAACCTTTCTGAAAGG
GCCAGATTGTAATATTCAAGACTGTGTGACCAAAAGGCCACATACAGTCTCTGTCATAACTACTCAACTCTGT
TTCTGAAGCAGGAAAGCCACACAGACAGTACATAAGGAATATGTGTCAGTGGGTTCCAGGCCAGACAAAACA
GATGGTGACCAAGACTTGGCCCTGGCTGAGTTGCTGACCCCTCATCTAAAAAATAGGCTATACTACAATTG
ACTTCCAGCACTTGAGAACGAGTTGAATACCAAGAATTATTCAATGGTCCCTCCAGTAACCTGCTAGAAACA
CAGAATTGGTCTGTATCTGACACTAGAACAAAATGAGGTTAAATAACATTGAATTAGAATCATAGAA
AACTGATTAGAAGAATACTTGATGTTATGATGATTGTTGAGTACAGATTTAAGTATGTTCTAAATATTGT
CTGCTGTAGTCTATTGCTGTATATGCTGAAATTGTTATGCTCATTTAGTATTGTTAGGAAATATT
TTCTAAGACCAGTTAGATGACTTATTCTGTAGTAATTCAATTGCTGTACCTGCTTGGTGGTTAGAAG
GAGGCTAGAAGATGAATTCAAGGACTTTCTCCAATAAAACATTGAGGTTAAATTGCTCATTCCCTTTGACAAGCTG
ACTGGATTCAATTAAACCATTTCATCAGTTCAATGGTAAATTCTGATTGATTGTTAAATGCGTTGG
AGAACTTGTATTAGGTAGTTACAGATTTATAAGGTGTTATATATTAGAAGCAATTATAATTACATCTG
TGATTCTGAACTAATGGTGTAAATTCAAGAGAAATGAAAGTGAAGATTCTCTGTTGTCATCGGCATTCC
AACTTTCTCTTGTGTTGTCAGTGTGCAATTGAAATATGCTGTTCTATAAATAATTGTTAAGAATAA

FIGURE 155

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48329
><subunit 1 of 1, 480 aa, 1 stop
><MW: 55240, pI: 9.30, NX(S/T): 2
MLFRNRFLLLALAALLAFVSLSLQFFHLIPVSTPKNGMSSKSRKRIMPDPVTEPPVTDPVY
EALLYCNIPSVAERSMEGHAPHFKLVSVHVFIRHGDRYPLYVIPKTKRPEIDCTLVANRKP
YHPKLEAFISHMSKGSGASFESPLNSLPLYPNHPLCEMGETQTVVQHLQNGQLLRDIYLK
KHKLLPNDWSADQLYLETTGKSRTLQSGLALLYGFLPDFWKIYFRHQPSALFCSGSCYCP
VRNQYLEKEQRRQYLLRLKNSQLEKTYGEMAKIVDVPTKQLRAANPIDSMLCHFCHNVSFPC
TRNGCVDMEHKVIKTHQIEDERERREKKLYFGYSLLGAHPILNQTIGRMQRATEGRKEELF
ALYSAHDVTLSALGLSEARFPFARLIFELWQDREKPSEHSVRLYNGVDVTFHTSF
CQDHHKRSPKPMCPLENLVRFVKRDMFVALGGSGTNYYDACHREGF
```

Signal sequence:

amino acids 1-18

FIGURE 156

AAAAAAGCTCACTAAAGTTCTATTAGAGCGAATACGGTAGATTCCATCCCTTTGAAGAACAGTACTGTGGA
GCTATTTAAGAGATAAAAACGAAATATCCTTCTGGAGTTCAAGATTGTGCAAGTAAATTGGTTAGGACTCTGAGC
GCCGCTGTTCACCAATCGGGGAGAGAAAAGCGGAGATCCTGCTCGCCTGCAACGCCCTGAAGCACAAAGCAGAT
AGCTAGGAATGAACCATCCCTGGGAGTATGTGAAACAACCGAGGAGCTCTGACTTCCAACGTCCCATTCTAT
GGGCGAAGGAACTGCTCTGACTTCAGTGGTTAAGGGCAGAATTGAAAATAATTCTGGAGGAAGATAAGA**ATGAT**
TCCCTGCGCAGTGCACCGGGACTACAAAGGGCTGTGCTGGGAATCCTCCTGGGGACTCTGTGGAGACCGG
ATGCACCCAGATACTGCTATTCAAGTCCGGAAGAGCTGGAGAAAGGCTCTAGGGTGGGAGACATCTCAGGGACCT
GGGGCTGGAGCCCGGGAGCTCGGGAGCGCGAGTCCGATCATCCCCAGGGTAGACCGGGAGGAGCTCTGTATGGGGCATCAAGTG
GAATCCGCGCAGCGCAGCTGGTCACGGGGCAGGATAGACCGGGAGGAGCTCTGTATGGGGCATCAAGTG
TCAATTAAATCTAGACATTCTGATGGAGGATAAGTAAAATATGGAGTAGAAGTAGAAGTAAGGGACATTAA
CGACAATGCGCCTTACTTCTGAAAGTGAATTAGAAAATTTAGTGAAGATGAGCAGCCACTGAGATGCGGTT
CCCTCTACCCCCACGCCCTGGATCCGGATATCGGGAAAGAACTCTGCAAGAGCTACGAGCTCAGGCCGAACACTCA
CTTCTCCCTCATCGTCAAAATGGAGCCAGGGTAGTAAGTACCCGAATTGGTGTGAAACGCCCTGGACCG
CGAAGAAAAGGCTGCTCACCACTGGCCTTACGCCCTCGACGGGGCGACCCGGTGCACAGGCACCGCG
CATCCGCGTGTGGTCTGGATGCGAACGACAACGACCCAGCGTTGCTCAGCCGAGTACCGCGAGCGTCC
GGAGAATCTGGCCTGGGACGCTGCTGTAGTCAACGCTACCGACCCCTGACGAAGGAGTCAATGCGGAAGT
GAGGTATTCTCTCCGGTATGTGGACGACAAGGCGGAGCAAGTTCAAACATAGATTGTAATTCAAGGACAATATC
AACAAATAGGGGAGTTGGACCACGAGGAGTCAGGATTCTACCAAGATGGAAGTGCAGCAATGGATAATGAGGATA
TTCTGCGCAGCCAAAGTCTGATCACTGTTCTGGACGTGAACGACAATGCCAGAAGTGGTCTCACCTCT
CGCCAGCTCGGTTCCGAAAACCTCTCCAGAGGGACATTAATTGCCCTTTAAATGTAATGACCAAGATTCTGA
GGAAAACGGACAGGTGATCTGTTCATCCAAGGAAATCTGCCCTTAAATTAGAAAATCTACGGAAATTACTA
TAGTTTAGTCACAGACATAGTCTGGATAGGGACACGGGCTCTAGCTACACATCACAGTACCGCCACTGACCG
GGGAACCCCGCCCTATCCACGGAAACTCATATCTCGCTGAACGTGGCAGACACCAACGACAACCCGGTCTT
CCCTCAGGCCCTTATTCGGTTATATCCAGAGAAACATCCAGAGGAGATTCCCTCGTCTCTGTGACCCG
CGACCCCGACTGTGAAGAGAACGCCAGATCACTTATCCCTGGCTGAGAACACCATCAAGGGCAAGCCTATC
GTCCTACGTGTCATCAACTCCGACACTGGGGTACTGTATGCGCTGAGCTCTCGACTACGAGCAGTCCGAGA
CTTGCAAGTGAAGTGTGGCGGGACAAACGGGACCCGCCCTCAGCAGCAACGTGTCGGTACGGTCCACTGGCGTGGAGCT
GCTGGACCAGAACGACAATGCGCCGAGATCCTGTACCCGCCCTCCCCACGGACGGTCCACTGGCGTGGAGCT
GGCTCCCCGCTCCGACAGGCCGGTACCTGGTACCAAGGTGGTGGCGGTGGACAGAGACTCCGGCAGAACGC
CTGGCTGTCTACCGCTGCTCAAGGCCAGCGAGCCGGACTCTCTCGGTGGTCTGCACACGGCGAGGTGCG
CACGGCGAGCCCTGCTGGACAGAGACCGCTCAAGCAGAGCCCTCGTAGTGGCGTCCAGGACACGCCAGCC
CCCTCTCTCCGCACTGTACGCTACCGTGGCGTGGCGACAGCATCCCCAAGTCTGGCGACCTGGCAG
CCTCGAGTCTCCAGCTAACTCTGAAACCTCAGACCTCACTCTGTACCTGGTGGTAGCGGTGGCGGTCTCCTG
CGTCTTCTGGCTTCTGTCATCTGCTGGCGCTCAGGCTGCCGCTGGACAAGTCACGCCCTGCTGCAGGC
TTCAGGAGGGGGCTTGACAGGAGCGCCGGCTGCACTTGTGGCGTGGACGGGGTGCAGGCTTCTGAGAC
CTATTCCCACGAGGTTCCCTCACCAACGGACTCGCGGAAGAGTCACCTGATCTCCCCAGCCCAACTATGAGA
CATGCTCGTCAGCCAGGAGAGCTTGAAAAAGCGAGCCCTTTGCTGTCAGGTGATTGGTATTTCTAAAGA
CAGTCATGGGTTAATTGAGGTGAGTTATATCAAATCTTCTTTCTTTTTAATTGCTCTGTCTCCCAAGC
TGGAGTGCAGCGGTACGATCATAGCTCACTGCCCTCAAACCTCTAGGCTCAAGCAATTATCCACCTTGCCT
CCGGTGTAAACAGGGACTACAGGTGCAAGGCCACCTACTGTCTGCCATCTATCTATCTATCTATCTATCT
CTATCTATCTATCTATCTATTACTTTCTGTCAGAGACGGGAGTCTACGCTGTAATCCAGTACTTGGAGGC
CGAGGCGGGTGGATCACCTGAGGGTTGGAGTTGAGACCGAGCCCTG**A**CCAACATGGAGAAACCCCGTCTATAACTAA
AAAAATACAAAATTAGCCGGCGTGGGTGTCATGTGTAATCCAGCTACTTGGAGGCTGAGTCAGGAGAAT
TGCTTAACCTGGGAGGTGGAGGTTGCAATGAGCTGAGATTGTGCCATTGCACTCCAGCCTGGCAACAAAGAGTG
AAACTCTATCTCA

FIGURE 157

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48306
><subunit 1 of 1, 916 aa, 1 stop
><MW: 100204, pI: 4.92, NX(S/T): 4
MIPARLHRDYKGLVLLGILLGTLWETGCTQIRYSVPEELEKGSRVGDISRDLGLEPRELAER
GVRIIPRGRTQLFALNPRSGSLVTAGRIDREELCMGAIKCQLNLDILMEDKVKIYGVEVEVR
DINDNAPYFRESELEIKISENAATEMRFPLPHAWDPDIGKNSLQSYELSPNTHFSLIVQNGA
DGSKYPELVLKRALDREEKAHHLVLTASDGGDPVRTGTARI RVMVLDANDNAPAFAQPEYR
ASVPENLALGTQLLVNATDPDEGVNAEVRYSFYVDDKAAQVFKLDNSGTISTIGELDHE
ESGFYQMEVQAMDNAGYSARAKVLITVLDVNDNAPEVVLTSASSVPENS PRGTLIALLNVN
DQDSEENGQVICFIQGNLPFKLEKSYGNYSLVTDIVLDREQVPSYNITVTATDRGTPPLST
ETHISLNVADTNNDNPPVFPQASYSAYIPENNPRGVSLVSVTAHDPCCEENAQITYSLAENTI
QGASLSSYVSINSDTGVLYALSSFDYEQFRDLQVKVMARDNGHPPRSSNVSLSFVLDQNDN
APEILYPALPTDGSTGVELAPRSAEPGYLVTKVVAVDRDSGQNAWLSYRLLKASEPGLFSVG
LHTGEVRTARALLDRDALKQSLVVAVQDHGQPPLSATVLTAVADSI PQVLADLGSLESPA
NSETSDLTLYLVAVAASCVFLAFVILLALRLRRWHKSRLLQASGGGLTGAPASHFVGVD
GVQAFLQTYSHESVSLTTDSRKSHLIFPQPQNYADMLVSQESFEKSEPLLLSGDSVFSKDSHGL
IEVSLYQIFFLFFFNCSVSQAGVQRYDHSSLRPQTPLKQLSHLCLRCNRDYRCKPPTVCLS
IYLSIYLSIYLSIYLLSCTDGSITPVI PVLWEAEAGGSPEVGSLRPA
```

Signal sequence:

amino acids 1-30

Transmembrane domains:

amino acids 693-711, 809-823, 869-888

FIGURE 158

CCCAGGCTCTAGTCAGGAGGAGAAGGAGGAGCAGGAGGTGGAGATTCCCAGTTAAAG
GCTCCAGAATCGTGTACCAGGCAGAGAACTGAAGTACTGGGGCCTCCTCCACTGGGTCCGAA
TCAGTAGGTGACCCCGCCCCTGGATTCTGGAAGACCTCACCATGGGACGCCCGACCTCGT
GCGGCCAAGACGTGGATGTTCTGCTCTGCTGGGGGAGCCTGGCAGGACACTCCAGGGC
ACAGGAGGACAAGGTGCTGGGGGTATGAGTGCCAACCCCATTGCAGCCTGGCAGGC
CCTTGTTCAGGCCAGCAACTACTCTGTGGCGGTGTCCTGTAGGTGGCAACTGGTCCTT
ACAGCTGCCACTGTAAAAAACGAAATACACAGTACGCCTGGAGACACAGCCTACAGAA
TAAAGATGGCCCAGAGCAAGAAATACCTGTGGTTAGTCCATCCCACACCCCTGCTACAACA
GCAGCGATGTGGAGGACCACAACCATGATCTGATGCTTCAACTGCGTGACCAGGCATCC
CTGGGTCAAAGTGAAGCCCATCAGCCTGGCAGATCATTGCACCCAGCCTGGCAGAAGTG
CACCGTCTCAGGCTGGGCACTGTACCAGTCCCCGAGAGAATTTCTGACACTCTCAACT
GTGCAGAAGTAAAAATCTTCCCCAGAAGAAGTGTGAGGATGCTTACCCGGGCAGATCACA
GATGGCATGGTCTGTGCAGGCAGCAGCAAAGGGCTGACACGTGCCAGGGCGATTCTGGAGG
CCCCCTGGTGTGTGATGGTGCACTCCAGGGCATCACATCCTGGGCTCAGACCCCTGTGGGA
GGTCCGACAAACCTGGCGTCTATACCAACATCTGCCGCTACCTGGACTGGATCAAGAAGATC
ATAGGCAGCAAGGGCTGATTCTAGGATAAGCACTAGATCTCCCTTAATAAACTCACAACTCT
CTGGTTC

FIGURE 159

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48336
<subunit 1 of 1, 260 aa, 1 stop
<MW: 28048, pI: 7.87, NX(S/T): 1
MGRPRPRAAKTWMFLLLGGAWAGHSRAQEDKVLGGHECQPHSQPWQAALFQQQLLCGGVL
VGGNWVLTAAHCKKPKYTVRLGDHSLQNKGPEQEIPVVQSI PHPCYNSSDVEDHNHDLMLL
QLRDQASLGSKVKPISLADHCTQPGQKCTVSGWGTVTSPRENFPDTLNCAEVKIFPQKKCED
AYPGQITDGMVCAGSSKGADTCQGDSGGPLVCDGALQGITSWGSDPCGRSDKPGVYTNICRY
LDWIKKIIGSKG
```

Important Features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 51-71

N-glycosylation site.

amino acids 110-113

Serine proteases, trypsin family, histidine active site.

amino acids 69-74 and 207-217

Tyrosine kinase phosphorylation site.

amino acids 182-188

Kringle domain proteins motif

amino acids 205-217

FIGURE 160

GGCGCCGGTGCACCGGGCGGGCTGAGCGCCTCCTGCGGCCCGGCCTGCGGCCCGGCCGC
CGCGCCGCCAACCCCCAACCCCCGGCCCGCGCCCCCTAGCCCCCGCCCGGGCCCGGCCGC
GCCCGCGCCAGGTGAGCGCTCCGCCCGCCCGAGGCCCCGCCCGGGCCGCCCGCCCCCG
CCCCGGCCGGCGGGGAACCGGGCGGATTCTCTCGCGCGTCAAACACCTGATCCCATAAAC
ATTCATCCTCCGGCGGCCCGCTGCGAGCGCCCCGCCAGTCCGCGCCGCCGCCCTCG
CCCTGTGCCCTGGCGGCCCTGGCACCAGGCCGGCCGAGGCCAGGCCAGAGCCGGCGGAGC
GGAGCGCGCCGAGCCTCGTCCCAGGCCGGCCGGGCCGGCCGTAGCGCGGCCCTGGA
TGCAGCCGGCCGCGGGGAGACGGCGCCGCCGAAACGACTTCACTCCGACGCC
CCCGCCAAACCCATCGATGAAGAGGGCGTCCGCTGGAGGGAGCCGGCTGCTGGCATGGTG
CTGTGGCTGCAGGCCTGGCAGGTGGCAGCCCCATGCCAGGTGCCTGCGTATGCTACAATGA
GCCAAGGTGACGACAAGCTGCCCGCAGCAGGGCCTGCAGGCTGTGCCGTGGCATCCCTG
CTGCCAGCCAGCGCATCTTCTGCACGGCAACCGCATCTCGCATGTGCCAGCTGCCAGCTTC
CGTGCCTGCCGAAACCTCACCATCTGTGGCTGCACTCGAATGTGCTGGCCGAATTGATGC
GGCTGCCTCACTGGCCTGGCCCTCTGGAGCAGCTGGACCTCAGCGATAATGCACAGCTCC
GGTCTGTGGACCCCTGCCACATTCCACGGCCTGGCCCTACACACGCTGCACCTGGACCGC
TGCAGGCCTGCAGGAGCTGGGCCGGGCTGTTCCGCGGCCCTGGCTGCCCTGCAGTACCTCTA
CCTGCAGGACAACCGCCTGCAGGCAGTGCCTGATGACACCTCCGCACCTGGCAACCTCA
CACACCTCTCCTGCACGGCAACCGCATCTCCAGCGTGCCAGCGCCCTTCCGTGGCTG
CACAGCCTCGACCGTCTCCTACTGCACCAAGACCGCGTGGCCCATGTGCACCCGATGCC
CCGTGACCTTGGCCGCTCATGACACTCTATCTGTTGCCAACATCTATCAGCGCTGCC
CTGAGGCCCTGGCCCCCTGCGTGCCCTGCAGTACCTGAGGCTCAACGACAACCCCTGGTG
TGTGACTGCCGGCACGCCACTCTGGCCTGGCTGCAGAAGTTCCGCGGCTCCTCC
GGTGCCTGCAGCCTCCGCAACGCCCTGGCTGGCGTGCACCTCAAACGCCAGTGC
ACCTGCAGGGCTGCCCTGAGGCCACCGCCCTTACCATCCATCTGAGCCGGCAGGGC
GATGAGGAGCCGCTGGGCTTCCAAGTGCTGCCAGCCAGATGCCGTGACAAGGCTCAGT
ACTGGAGCCTGGAAGACCAAGCTTCCGGCAGGAATGCCGTGAGGGACCGCGTGC
ACAGCCCCGGCAACGCCCTGGCCACGGCACATCAATGACTCACCC
CCTGGCTCTGCTGAGCCCCGCTCACTGCAGTGCAGGCCAGGGCTCGAGCC
CCCCACCTCGGGCCCTGCCGGAGGCCAGGCTGTTACGCAAGAACCGC
GCCGTCTGGCCAGGCAGGCAGCGGGGTGGCGGACTGGTACTCAGAAGGCT
CTACCCAGCCTCACCTGCAGCCTACCCCCCTGGCCTGGCGCTGGTGC
TGGGCCCTGCTTGAACCCCCAGCGGACACAAGAGCGTGC
GGGGTCTCTCCACGCCGCAAGCCAGGCCGGCGGCCACCGTGGGGCAGGCC
GTCCTCCCTGATGGAGCCCTGCCGCCACCCCATCTCCACCC
TTCGGCGGCAGCGTTGTTCCAGAACGCCCTCCACCCAGATGCC
GCATTTATTTACTTGTGTAAAATATCGGACGACGTGGAATAAGAGCT
AAAA

FIGURE 161

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44184
><subunit 1 of 1, 473 aa, 1 stop
><MW: 50708, pI: 9.28, NX(S/T): 6
MKRASAGGSRLLAWVLWLQAWQVAAPCPGACVCYNEPKVTTSCPQQGLQAVPGIPAASQRI
FLHGNRISHVPAASFACRNLTILWLHSNVLARIDAAAFTGLALLEQLDLSDNAQLRSVDPA
TFHGLGRLHTLHLDRCGLQELGPGLFRGLAALQYLYLQDNALQALPDDTFRDLGNLTHLFLH
GNRISSVPERAFRGLHSDLRLLLHQNRVAHVPHAFRDLGRLMTLYLFANNLSALPTEALAP
LRALQYLRLLNDNPWVCDCRARPLWAWLQKFRGSSSEVPCSLPQRLAGRDLKRLAANDLQGCA
VATGPYHPIWTGRATDEEPLGLPKCCQPDAADKASVLEPGRPASAGNALKGRVPPGDSPPGN
GSGPRHINDSPFGTLPGSAEPLTAVRPEGSEPPGFPTSGPRRPGCSRKNRTRSHCRLGQA
GSGGGGTGDSESGALPSLTCSLTPLGLALVLWTVLGPC
```

Important features:

Signal peptide:

amino acids 1-26

Leucine zipper pattern.

amino acids 135-156

Glycosaminoglycan attachment site.

amino acids 436-439

N-glycosylation site.

amino acids 82-85, 179-183, 237-240, 372-375 and 423-426

vWFC domain

amino acids 411-425

FIGURE 162

GGAAGTCCACGGGAGCTGGATGCCAAGGGAGGACGGCTGGTCCTCTGGAGAGGACTAC
TCACTGGCATATTCTGAGGTATCTGTAGAATAACCACAGCCTCAGATACTGGGGACTTAC
AGTCCCACAGAACCGTCCTCCCAGGAAGCTGAATCCAGCAAGAACAAATGGAGGCCAGCGGA
AGCTCATTGCAGACAAAGGCAAGTCCTTTCTCCTTTGAGGCTATCTCTGGCG
GGCGCGCGGAACCTAGAAGCTATTCTGTGGAGGAAACTGAGGGCAGCTCCTTGTAC
CAATTAGCAAAGGACCTGGCTGGAGCAGAGGAATTCTCCAGGCGGGGGTAGGGTTG
TTTCCAGAGGAACAAACTACATTGCAGCTCAATCAGGAGACCGCGATTGTTGCTAAAT
GAGAAATTGGACCGTGAGGATCTGTGCGGTACACAGAGCCCTGTGTACGTTCAA
GTTGCTAGAGAGTCCCTCGAGTTTCAAGCTGAGCTGCAAGTAATAGACATAAACGACC
ACTCTCCAGTATTCTGGACAAACAAATGTTGGTAAAGTATCAGAGAGCAGTCCTCCTGGG
ACTACGTTCTCTGAAGAACGCGAAGACTTAGATGTAGGCCAAACAAATATTGAGAACTA
TATAATCAGCCCCAACTCCTATTTCGGGTCTCACCCGCAAACGCAGTGATGGCAGGAAAT
ACCCAGAGCTGGTCTGGACAAAGCGCTGGACCGAGAGGAAGAAGCTGAGCTCAGGTTAAC
CTCACAGCACTGGATGGTCTCCGCCAGATCTGGACTGCTCAGGTTACATCGAAGT
CCTGGATGTCAACGATAATGCCCTGAATTGAGCAGCCTTCTATAGAGTGCAGATCTG
AGGACAGTCCGGTAGGCTCTGGTTGTGAAGGTCTTGCCACGGATGTAGACACAGGAGTC
AACGGAGAGATTCTTCTATTCACTTTCCAAGCTTCAGAAGAGATTGGCAAAACCTTAAGAT
CAATCCCTTGACAGGAGAAATTGAACCTAAAAAAACAACCTCGATTTGAAAAACTTCAGTCCT
ATGAAGTCAATATTGAGGCAAGAGATGCTGGAACCTTTCTGGAAAATGCACCGTTCTGATT
CAAGTGATAGATGTGAACGACCATGCCAGAAGTTACCATGTCTGCATTACAGCCAAAT
ACCTGAGAACGCGCTGAAACTGTGGTTGCACTTTCAGTGTTCAGATCTGATTCAAGGAG
AAAATGGAAAATTAGTTGCTCCATTAGGAGGATCTACCCCTCTCTGAAATCCGCGGAA
AACTTTACACCTACTAACGGAGAGACCAGACAGAGAAAGCAGAGCGGAATAACACAT
CACTATCACTGTCACTGACTTGGGACCCCTATGCTGATAACACAGCTCAATATGACCGTGC
TGATCGCCGATGTCATGACAACGCTCCGCCCTCACCCAAACCTCCTACACCCCTGTC
CGCGAGAACAAACAGCCCCGCCCTGCACATCCGAGCGTCAGCGCTACAGACAGAGACTCAGG
CACCAACGCCAGGTACCTACTCGCTGCTGCCGCCAGGACCCGACCTGCCCTCACAT
CCCTGGTCTCCATCAACGCGGACAACGCCACCTGTTGCCCTCAGGTCTCTGGACTACGAG
GCCCTGCAGGGTCCAGTCCGCTGGCGCTTCAGACCACGGCTCCCGCGCTGAGCAG
CGAGGCCTGGTGCCTGGTGGTGGTGGACGCCAACGACAACCTGCCCTCGTGTGCTTAC
CGCTGCAGAACGGCTCCGCCCTGCACCGAGCTGGTGCCCGGGCGAGCCGGCTAC
CTGGTGACCAAGGTGGTGGCGGTGGACGGCGACTCGGGCCAGAACGCTGGCTGTC
GCTGCTCAAGGCCACGGAGCTCGGTCTGTCGGCGTGTGGCGCACAAATGGCGAGGTGCG
CCGCCAGGCTGCTGAGCGAGCGCACGCCAGCAGCTGACGTGCTCTGGTGGACGGCTTC
AATGGCGAGCCTCCGCGCTCGGCCACGCCAGCAGCTGACGTGCTCTGGTGGACGGCTTC
CCAGCCCTACCTGCCCTCCGGAGGCGGCCAGCAGGCCAGGCGACTTGCTCACCG
TCTACCTGGTGGTGGCGTTGGCCTCGGTGTCCTCGCTCTTCTGGTGTGCTCTGGT
GTGGCGGTGCGCTGTGAGGAGGAGCAGGGCGGCCCTCGGTGGTGTGCTGCTGGT
GGGCCCTCCAGGGCATCTGTGGACATGAGCGGCAGCAGGACCCATCCCAGAGCTACC
AGTATGAGGTGTCTGGCAGGAGGCTCAGGGACCAATGAGTTCAAGTTCTGAAGCCGATT
ATCCCCAACTTCCCTCCCCAGTGCCCTGGAAAGAAATACAAGGAAATTCTACCTCCCCAA
TAACCTGGGTTCAATATTCAAGTGACCCATAGTTGACTTTACATTCCATAGGTATTT
TGTGGCATTCCATGCCAATGTTATTCCCCAATTGTGTATGTAATATTGTACGGAT
TTACTCTGATTTCTCATGTTCTTCTCCCTTGTTAAAGTGAACATTACCTTATT
CCTGGTTCTT

FIGURE 163

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48314
<subunit 1 of 1, 798 aa, 1 stop
<MW: 87552, pi: 4.84, NX(S/T): 5
MEASGKLICRQRQVLFSFLLGLSLAGAAEPRSYSVVEETEGSSFVTNLAKDLGLEQREFSR
RGVRVVSRGNKLHLQLNQETADLLLNEKLDREDLCGHTEPVCVLFQVLLESPEFFQAELOV
IDINDHSPVFLDKQMLVKVSESSPPGTTFPLKNAEDLDVGQNNIENYIISPNSYFRVLTRKR
SDGRKYPELVLKDALKALDREEEAEELRLTALDGGSPPRSGTAQVYIEVLDVNDNAPEFEQPFY
RVQISEDSPVGFLVVKVSATDVDTGVNGEISYSLFQASEEIGKTFKINPLTGEIELKKQLDF
EKLQSYEVNIEARDAGTFSGKCTVLIQVIDVNDHAPEVTMSAFTSPIPENAPETVVALFSVS
DLDSENGKISCSIQEDLPFLLKSAENFYTLLTERPLDRESRAEYNITITVTDLGTPMILITQ
LNMTVLIADVNDNAPAFQTQSYTLFVRENNSPALHIRSVSATDRDSGTNAQVTYSLLPPQDP
HLPLTSLVSINADNGHLFALRSLDYEALQGFQFRVGASDHGSPALSSEALVRVVVLDANDNS
PFVLYPLQNGSAPCTELVPRAAEPGYLVTKVVAVDGDSGQNAWLSYQLLKATELGLFGVWAH
NGEVRTARLLSERDAAKHRLVVLVKDNGEPPRSATATLHVLLVDGFSQPYLPLPEAAPTQAQ
ADLLTVYLVVALASVSSLFLFSVLLFVAVRLCRRSRAASVGRCLVPEGPLPGHLVDMMSGRT
LSQSYQYEVCLAGGSGTNEFKFLKPIIPNFPQCPGKEIQGNSTFPNNFGFNIQ
```

Important features:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 685-712

Cadherins extracellular repeated domain signature.

amino acids 122-132, 231-241, 336-346, 439-449 and 549-559

ATP/GTP-binding site motif A (P-loop).

amino acids 285-292

N-glycosylation site.

amino acids 418-421, 436-439, 567-570 and 786-789

FIGURE 164

ACCCACCGCGTCCGCCACCGCGTCCGCCACCGCGTCCGCCACCGCGTCCGCCGTAGCCGTGC
GCCGATTGCCTCTGGCCTGGCAATGGTCCCGGCTGCCGGTCAACGACCGCCCCCGCGTCAT
GCGGCTCCTCGGCTGGCAAGTATTGCTGTGGTGTGGACTTCCCGTCCGCCGTGG
AGGTTGCAGAGGAAAGTGGTCGCTTATGGTCAGAGGAGCAGCCTGCTCACCTCTCCAGGTG
GGGGCTGTGTACCTGGGTGAGGAGGAGCCTGCATGACCCGATGGCCAGGACAGGGCAGC
AGAAGAGGCCAATCGGGTGCTGGGCTGGACACCCAAGGCGATCACATGGTATGCTGTCTG
TGATTCCCTGGGAAGCTGAGGACAAAGTGAGTTAGAGCCTAGCGCGTCACCTGTGGTGT
GGAGGAGCGGAGGACTCAAGGTCAACGTCCGAGAGAGCCTTCTCTGGATGGCGCTGG
AGCACACTCCCTGACAGAGAAGAGGAGTATTACACAGAGCCAGAAGTGGCGGAATCTGACCG
CAGCCCCGACAGAGGACTCCAATAACACTGAAAGTCTGAAATCCCCAAGGTGAACGTGAG
GAGAGAAACATTACAGGATTAGAAAATTCACTCTGAAAATTAAATATGTCACAGGACCT
TATGGATTTCTGAACCCAAACGGTAGTGACTGTACTCTAGCCTGTTACACCCCGTGGT
GCCGCTTTCTGCCAGTTGGCCCTCACTTAACCTCTGCCCGGGCATTCCAGCTCTT
CACTTTGGCACTGGATGCATCTCAGCACAGCAGCCTTCTACCAGGTTGGCACCGTAGC
TGTTCTAATATTATTCAAGGAGCTAACCAATGCCAGATTAAATCATACAGATC
GAACACTGGAAACACTGAAATCTCATTAAATCAGACAGGTATAGAAGCCAAGAAGAAT
GTGGTGGTAACTCAAGCCGACCAATAGGCCCTTCCCAGCACTTGATAAAAAGTGTGGA
CTGGTTGCTTGTATTTCTTATTCTTTAATTAGTTTATTATGTATGCTACCATTGAA
CTGAGAGTATTCGGTGGCTAATTCCAGGACAAGAGCAGGAACATGTGGAGTAGTGATGGTCT
GAAAGAAGTTGGAAAGAGGAACCTCAATCCTCGTTCAGAAATTAGTGCTACAGTTCTA
CATTTCTCCAGTGACGTGTTGACTTGAAACTCAGGCAGGATTAAAGAATCATTGTTGAA
CAACTGAATGTATAAAAAATTATAAACTGGTGTAACTAGTATTGCAATAAGCAAATGC
AAAAATATTCAATAG

FIGURE 165

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48333
><subunit 1 of 1, 360 aa, 1 stop
><MW: 39885, pI: 4.79, NX(S/T): 7
MVPAAGRRPPRVMRLLGWWQVLLWVLGLPVRGVEVAEESGRLWSEEQPAHPLQVGAVYLGE
ELLHDPMGQDRAAEEANAVLGLDTQGDHMVMLSVPGEAEDKVSSEPSGVTGAGGAEDSRC
NVRESLFSLDGAGAHFPDREEYYTEPEVAESDAAPTEDSNNTESLKSPKVNCEERNITGLE
NFTLKILNMSQDLMDFLNPNQSDCTLVLFYTPWCRCFSASLAPHFNSLPRAFPALHFLALDAS
QHSSLSTRFGTVAVPNILLFQGAKPMARFNHTDRTLETLKIFIFNQTGIEAKKNVVVTQADQ
IGPLPSTLIKSVDLLVFSLLFISFIMYATIRTESIRWLIPGQEQEHEVE
```

Important features:

Signal peptide:

amino acids 1-25

Transmembrane domain:

amino acids 321-340

Homologous region to disulfide isomerase

amino acids 212-302

N-glycosylation site.

amino acids 165-168, 181-184, 187-190, 194-197, 206-209, 278-281
and 293-296

Thioredoxin domain

amino acids 211-227

FIGURE 166

CCCGGCTCCGCTCCCTCTGCCCTCGGGGTCGCGCGCCCACGATGCTGCAGGGCCCTGGCT
CGCTGCTGCTGCTCTTCCTCGCCTCGCACTGCTGCCTGGCTCGCGCGGGCTTTCCCTC
TTTGGCCAGCCCACCTCTCCTACAAGCGCAGCAATTGCAAGCCCACCTGGTCAACCTGCA
GCTGTGCCACGGCATCGAATACCAGAACATGCGGCTGCCAACCTGCTGGCACGAGACCA
TGAAGGAGGTGCTGGAGCAGGCCGGCGCTGGATCCGCTGGTCAAGCAGTGCACCCG
GACACCAAGAAGTCCGTGCTCGCTCTCGCCCCGTGCCTCGATGACCTAGACGAGAC
CATCCAGCCATGCCACTCGCTCGGTGAGGTGAAGGACCGCTGCACCCGGTCAATGTCG
CCTTCGGCTTCCCTGGCCCGACATGCTTGAGTGCACCGTTCCCCCAGGACAACGACCTT
TGCATCCCCCTCGCTAGCAGCGACACCTCTGCCAGCCACCGAGGAAGCTCAAAGGTATG
TGAAGCCTGCAAAATAAAATGATGATGACAACGACATAATGGAAACGCTTGAAATG
ATTTGCACTGAAAATAAAAGTGAAGGAGATAACCTACATCAACCGAGATAACAAATCATC
CTGGAGACCAAGAGCAAGACCATTACAAGCTGAACGGTGTCCGAAAGGGACCTGAAGAA
ATCGGTGCTGTGGCTCAAAGACAGCTTGCACTGTGAGGAGATGAACGACATCAACG
CGCCCTATCTGGTCATGGACAGAAACAGGGTGGGAGCTGGTATCACCTCGGTGAAGCGG
TGGCAGAAGGGCAGAGAGAGTTCAAGCGCATCTCCCGCAGCATCCGCAAGCTGCAGTGCTA
GTCCCCGGCATCCTGATGGCTCCGACAGGCCCTGCTCCAGAGCACGGCTGACCATTCTGCTCC
GGGATCTCAGCTCCGTTCCCAAGCACACTCCTAGCTGCTCCAGTCTCAGCCTGGCAGCT
TCCCCCTGCCTTGCACGTTGCATCCCCAGCATTCTGAGTTATAAGGCCACAGGAGTG
GATAGCTGTTTACCTAAAGAAAAGCCCACCGAATCTTGTAGAAATATTCAAACATAATA
AAATCATGAATATTTAA

FIGURE 167

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50920
><subunit 1 of 1, 295 aa, 1 stop
><MW: 33518, pI: 7.74, NX(S/T): 0
MLQGPGSLLLLFLASHCCLGSARGLFLFGQPDFSYKRSNCKPIPVNLQLCHGIEYQNMRLPN
LLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLCSLFAPVCLDDLDETIQPCHSLCVQVKDR
CAPVMSAFGFPWPDMLECDRFPQDNDLCIPLASSDHLLPATEEAPKVEACKKNNDNDIM
ETLCKNDFAALKIKVKEITYINRDTKIILETKSCTIYKLNGVSERDLKKSVLWLKDSLQCTCE
EMNDINAPYLVMGQKQGGELVITSVKRWQKGQREFKRISRSIRKLQC
```

Important features:

Signal peptide:

amino acids 1-20

Cysteine rich domain, homologous to frizzled N terminus

amino acids 6-153

FIGURE 168

GTGGAGGCCGCGACGATGGCGGGCGACGGAGGCCGAGACGGGTTGGCGAGCCCCGGG
CCCTGTGCGCGCAGGGGCCACCGCACCTACGCGCGCGCTGGGTGTTCTGCTCGCGATC
AGCCTGCTCAACTGCTCCAACGCCACGCTGTGGCTCAGCTTGCACCTGTGGCTGACGTCA
TGCTGAGGACTTGGTCCTGTCCATGGAGCAGATCAACTGGCTGTCACTGGTCTACCTCGTGG
TATCCACCCATTGGCGTGGCGGCATCTGGATCCTGGACTCCGTGGCTCCGTGCGCG
ACCATCCTGGGTGCGTGGCTGAACCTTGCCGGAGTGTGCTACGCATGGTGCCTGCATGGT
TGTTGGGACCCAAAACCCATTGCGCTTCATGGGTGGCCAGAGCCTCTGTGCCCTGCC
AGAGCCTGGTCATCTCTCCAGCCAAGCTGGCTGCCTGTGGTCCCAGAGCACCAGCGA
GCCACGGCCAACATGCTCGCCACCATGTCGAACCTCTGGCGTCCTGTGGCAATGTGCT
GTCCCCGTGCTGGTCAAGAAGGGTGAGGACATTCCGTTAATGCTCGGTGTCTATACCATCC
CTGCTGGCGTCTGCCTGCTGTCCACCATCTGCCTGTGGAGAGTGTGCCCCCCCC
CCCTCTGCCGGGCTGCCAGCTCACCTCAGAGAAGTCCCTGGATGGCTCAAGCTGCAGCT
CATGTGGAACAAGGCCTATGTCATCCTGGCTGTGCTGGGGGAATGATGGGATCTCTG
CCAGCTTCTCAGCCCTCTGGAGCAGATCCTCTGTGCAAGCGGCCACTCCAGTGGTTTCC
GGCCTCTGTGGCGCTCTTCATCACGTTGGATCCTGGGGCACTGGCTCTCGGCCCTA
TGTGGACGGACCAAGCACTTCACTGAGGCCACCAAGATTGGCCTGTGCCTGTTCTCTGG
CCTGCGTGCCTTGGCTGGTCCAGCTGCAGGGACAGACCCCTGCCCTGGCTGCCACC
TGCTCGCTGCTCGGGCTGTTGGCTTCGGTGGGCCATGGAGTTGGCGGTGCA
GTGTTCCCTCCCCGTGGGGAGGGGGCTGCCACAGGCATGATCTTGCTGGGAGGCCG
AGGAAATACTCATCATGCTGGCAATGACGGCACTGACTGTGCAGCCTGGAGGCCCTTG
TCCACCTGCCAGCAGGGGAGGATCCACTGACTGGACAGTGTCTCTGCTGATGGCCGG
CCTGTGCACCTCTTCAGCTGCATCCTGGCGCTTCTTCCACACCCATACGGCGCCTGC
AGGCCGAGTCTGGGAGCCCCCTCACCGTAACGCCGTGGCGGCCAGACTCAGGCCG
GGTGTGGACCGAGGGGGAGCAGGAAGGGCTGGGTCTGGGGCCAGCACGGCACTCCGGA
GTGCACGGCGAGGGGGGCTCGCTAGAGGACCCAGAGGGCCGGAGCCCCACCCAGCCT
GCCACCGAGCGACTCCCCGTGCGCAAGGCCAGCAGCCACCGACGCCCTCCGCCCGC
AGACTCGCAGGCAGGGTCCAAGCGTCCAGGTTATTGACCCGGCTGGGTCTCACTCCTC
CTCCTCCCCGTGGGTGATCACGTAGCTGAGCGCCTGTAGTCCAGGTTGCCGCCACATCGA
TGGAGGCCGAACATCTGGTCCACCTGCGGGCGGGCGAAAGGGCTCCTGCGGGCT
CCGGGAGCGAATTACAAGCGCGACCTGAAAA

FIGURE 169

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50988
><subunit 1 of 1, 560 aa, 1 stop
><MW: 58427, pI: 6.86, NX(S/T): 2
MAGPTEAETGLAEPRALCAQRGHRTYARRWFLLAISLLNCSNATLWLSFAPVADVIAEDLV
LSMEQINWLSLVYLVVSTPFGVAAIWILDVGGLRAATILGAWLNFAGSVLRMVPVPCMVGQTQN
PFAFLMGGQSLCALAQSLVIFSPAKLAALWFPEHQRATANMLATMSNPLGVLVANVLSPVLV
KKGEDIPLMLGVYTIPAGVVCLLSTICLWESVPPTPPSAGAASSTSEKFLDGLKLQLMWNK
YVILAVCLGGMIGISASFSALEQILCASGHSSGFSGLCGALFITFGILGALALGPYVDRTK
HFTEATKIGLCLFSLACVPFALVSQLQGQTLALAATCSLLGLFGFSVGPVAMELAVECSFPV
GEGAATGMI FVLGQAEGILIMLAMTALTVRSEPSLSTCQQGEDPLDWTVSLLL MAGLCTFF
SCILAVFFHTPYRRLQAESGEPPSTRNAVGGA DSGPGVDRGGAGRAGVLPSTATPECTARG
ASLEDPRGPGSPHPACHRATPRAQGPAATDAPS RPGRLAGRQASRFIDPAGSHSSFSSPWVIT
```

Important features:

Signal peptide:

amino acids 1-44

Transmembrane domains:

amino acids 61-79, 98-112, 126-146, 169-182, 201-215, 248-268,
280-300, 318-337, 341-357, 375-387, 420-441

N-glycosylation site.

amino acids 40-43 and 43-46

Glycosaminoglycan attachment site.

amino acids 468-471

FIGURE 170

GTCCCACATCCTGCTCAACTGGGTCAAGGTCCTTCTTACCTGGACAGCTCTTGTCCATCTTGCTGAAGTGGACCAAC
TAGTTCCCCAGTAGGGGGTCTCCCCTGGCAATTCTTGATCGGCGTTGGACATCTCAGATCGCTCCAATGAAGA
TGGCTTGCCTTGGGGCTGCTTGTTCATAATCATCTAATCTACTATGGGACAAGGTTGTGCCGGCAGCTCTGGGG
AAGGAGCACGGGCTGATCAAGGCATCCAGGAAACACTGGAGGACTGTCCAGCCTGAAAGAACTCTAGTGGTT
TCTGAATCTAGCCCACTTGGCGGTAAAGCATGATGCAACTTCTGCAACTTCTGCTGGGGCTTTGGGGCAGGTGG
CTACTTATTCTTTAGGGATTGTCAAGGAGGTGACCACTCTCACGGTAAATAACCAAGTGTCAAGAGGAAGTGC
ATCTGGTACAGTGTCAAGGTCGGAGCTCCAGGAACCTGGCCGGAGGAGAGGGCGGAGGCAAGCTGGGGCCGCCTT
CCAGGTGTTGCAAGCTGCCTCAGGCCTCCCCATTAGGTGACTCTGAGGAAGGCTTGCTCAGCACAGGCAGGC
GCTGGATCGAGAGCAGCTGTGCCAGTGGGATCCTGCTGGTTCTTGATGTGCTGCCACAGGGATT
GGCTCTGATCCATGTGGAGATCCAAGTGCTGGACATCAATGACCACCCAGGTTCCAAAGGCGAGCAGGA
GCTGGAAATCTCTGAGAGCGCCTCTGCGAACCCGGATCCCCCTGGACAGAGCTTGAACACAGGCC
TAACACCCCTGCACACCTACACTCTGTCCTCCAGTGAGCACTTGCCTGGATGTCATTGTCAGGCTGAGAC
CAAACATGCAGAACTCATAGTGGTAAGGAGCTGGACAGGGAAATCCATTCACTTTGATCTGGTGTAACTGC
CTATGACAATGGGAAACCCCCCAAGTCAGTACCAAGCTTGGTCAAGGTCAACGTCTGGACTCCAATGACAATAG
CCCTGCGTTGCTGAGAGTTCACTGGCACTGGAAATCCAAGAAGATGCTGCACCTGGTACGCTCTCATAAAACT
GACCGCCACAGACCCCTGACCAAGGCCAATGGGAGGTGGAGTTCTTCAGTAAGCACATGCCTCAGAGGT
GCTGGACACCTTCAGTATTGATGCCAAGACAGGCCAGGTCAATTCTGCGTCACCTCTAGACTATGAAAAGAACCC
TGCCCTACGAGGTGGATGTTCAAGGCAAGGGACCTGGTCCCAATCCTATCCAGCCATTGCAAAGTTCTCATCAA
GGTTCTGGATGTCAATGACAACATCCCAAGCATTCCACGTACATGGCCTCCAGCCACTGGTGTCAAGAAGC
TCTTCCAAGGACAGTTTATTGCTCTTGTCACTGGCAGATGACTTGGATTCAAGGACACAATGGTTGGTCAACTG
CTGGCTGAGCCAAGAGCTGGGCAACTTCAGGCTGAAAGAAACTAATGGCAACACATACATGTTGCTAACCCTGC
CACACTGGACAGAGGAGCAGTGGGCAAAATACCCCTACTCTGTTAGCCCAGGAAAGACTCCAGCCCTTATC
AGCCAAGAAACAGCTCAGCATTCAAGTCACTGGACATCAACGACAATGCACTGTGTTGAGAAAAGCAGGTATGA
AGTCTCCACGGGGAAAACAACCTTACCCCTCTTCACCTCATTACCATCAAGGCTCATGATGCAAGACTTGGGCAT
TAATGGAAAAGTCTCATACCGCATCCAGGACTCCCCAGTTGCTCACTTAGTAGCTATTGACTCCAACACAGGAGA
GGTCACTGCTCAGAGGTCACTGAACATGAAGAGATGGCCGGCTTGAGTTCAGGTGATGCAAGGACAGCGG
GCAACCCATGCTGCATCCAGTGTCTGTGGTCAGCCTCTGGATGCCAATGATAATGCCCAAGGGTGGT
CCAGCCTGCTCAGCGATGGAAAAGCCAGGCCACTGACACACCTCCACTGGCAACTCACAGCTCCGGCATT
CCTTTGACAACCATTGTGGCAAGAGATGCAACTCAGGGGCAATGGAGAGGCCCTTACAGCATCCGCAATGG
AAATGAAGGCCACCTCTCATCCTCAACCCCTACAGGGGCACTGTCATGTCACCAATGCCAGCAGCCT
CATTGGAGTGAGTGGGAGCTGGAGATAGTAGTAGAGGACCAAGGAAGCCCCCCTTACAGACCCGAGCCCTGTT
GAGGGCATGTTGTCACCAAGTGTGGACCACTGGAGGGACTCAGCCCGAAGCCTGGGGCTTGAGCATGTCAT
GCTGACGGTGTACTGCTGGCTGACTGTTGGCATCTCGGGTTGATCTGGCTTGTTCATGTCATCTGCCG
GACAGAAAAGAAGGACAACAGGCCTACAACACTGTCGGAGGCCAGTCCACCTACCAGCAGGCCAAAGAGGCC
CCAGAAACACATTCAAGAAGGCAGACATCCACCTCGTGCCTGTGTCAGGGTCAGGCAAGGTGAGCCTTGTGAAGT
CGGGCAGTCCCACAAAGATGTGGACAAGGAGGCGATGATGGAAGCAGGCTGGGACCCCTGCTGCAGGCCCTT
CCACCTCACCCGACCTGTACAGGACGCTGCGTAATCAAGGCAACCAGGGAGCAGCCGGAGAGCCGAGAGGT
GCTGCAAGACACGGTCAACCTCTTTCACCATCCAGGCAGAGGAATGCTCCGGGAGAAACCTGAACCTTCC
CGAGCCCCAGCCTGCCACAGGCCAGCCACGTTCCAGGCCCTGTGAAGGTTGCAAGGCGAGGCCACAGGGAGGCTGGC
TGGAGACCGGGCACTGGAGGAAGCCCCACAGAGGCCACAGGCCCTCTGCAACCCCTGAGACGGCAGGCCACATCT
CAATGGCAAAGTGTCCCTGAGAAAAGAATCAGGGCCCGTCAGATCTCGGGCTGGCTGGCTGTCTGGC
TGCCTTCGCGGAGCGGAACCCCGTGGAGGAGCTCACTGTTGATTCTCCTCTGTCAGCAAATCTCCAGCTGCT
GTCCTTGCTGCATCAGGGCAATCCAGCCAAACCAACCCAGGAGGAATAAGTACTTGGCAAGCCAGGAGG
CAGCAGGAGTGAATCCCAGACACAGATGCCCAAGTGCAAGGGCTGGAGGCCAGACAGACCCAGAACAGGAGGA
AGGGCTTGGATCCTGAAAGAGGACCTCTGTGAAGCAACTGCTAGAAGAAGAGCTGTCAAGTCTGCTGGACCC
CAGCACAGGTCTGGCCCTGGACCCGCTGAGGCCCTGACCCGGCTGGATGGCGAGACTCTTGGCCCTCAC
CACCAACTACCGTGACAATGTGATCTCCCCGGATGCTGCAGCCACGGAGGCCAGGACCTTCCAGCAGTTCGG
CAAGGCAGAGGCCACAGAGCTGAGCCAAACAGGCACGCCAGGCTGCCAGCACCTTGCTCGGAGATGAGCTCACT
GCTGGAGATGCTGCTGGAACAGCGCTCCAGCATGCCGTGGAGGCCCTCCAGGGCGCTGCCGGCTCTCGGT
CTGGGGAGGACCCCTCAGTTAGACTTGGCCACCGTCAGGCTCAGGCATGAAAGTGTCAAGGGGACCCAGGTGG
AAAGACGGGAGCTGAGGGCAAGAGCAGAGGCCAGCAGCAGCAGCAGGCCAGGTTGTGAGATAAGTGT
CTGGATCCAAGAACAGGGGCTGAGGATCTGTGGACAAGAGCTGGTTCTAAAATCTGTAACACTAGCTAG
CGCGGGCTGAGAACTTAAAGGGTCACTGATGCTACCCCCACAGAGGCCAGGACTAACAGCTGAC
TGACCAAAGCAGCCCTGTAAAGCAGCTGAGTCTTGGAGGACAGGGACGGTGTGGCTGAGATAAGTGT
TCCTGGCAAAACATATGTGGAGCACAAGGGTCAGTCCCTGGCAGAACAGATGCCACGGAGTATCACAGGCCAG
AAAGGGTGGCCTTCTGGGTAGCAGGAGTCAGGGGCTGTACCTGGGGTGCACAGGAAATGCTCTGACCTAT
CAATAAAGGAAAAGCAGTAAAAAAAAAAAAAA

FIGURE 171

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48331
<subunit 1 of 1, 1184 aa, 1 stop
<MW: 129022, pI: 5.20, NX(S/T): 5
MMQLLQLLLGPGGYLFLGDCQEVTTLTVKYQVSEEVPSGTIVGKLSQELGREERRQA
GAAFQVLQLPQALPIQVDSEEGLLSTGRRLDREQLCRQWDPCCLVSFDVLATGDLALIHEIQ
VLDINDHQPRFPKGEGELEISESASLRTRIPLDRALDPDTGPNTLHTYTLSPSEHFALDVIV
GPDETAKHAELIVVKELDREIHSFFDLVLTAYDNGNPPKSGTSLVKVNVLDSNDNSPAFAESS
LALEIQEDAAPGTLLIKLTATDPDQGPNGEVEFFLSKHMPEVLDTFSIDAKTGQVILRRPL
DYEKNPAYEVDVQARDLGPNPPIPAHCKVLIKVLVDVNDNIPSIHVTWASQPSLVSEALPKDSF
IALVMADDLDSGHNGLVHWCWLSQELGHFRLKRTNGNTYMLLTNATLDREQWPKYTLTLAQN
QGLQPLSAKKQLSIQISDINDNAPVFEKSRYEVSTRENNLPSLHLITIKAHDADLGINQKVS
YRIQDSPVAHLVAIDSNTGEVTAQRSLNYEEMAGFEFQVIAEDSGQPMЛАSSVSVWVSSLDA
NDNAPEVVQPVLSDGKASLSVLVNASTGHLLVPIETPNGLGPAGTDTPPLATHSSRPFLTT
IVARDADSGANGEPLYSIRNGNEAHLFILNPHTGQLFVNVTNASSLIGSEWELEIVVEDQGS
PPLQTRALLRVMFVTSVDHLRDSARKPGALSMSMLTVICLAVLLGIFGLILALFMSICRTEK
KDNRAYNCREAESTYRQQPKRPQKHIQKADIHLVPVLRGQAGEPCEVGQSHKDVDKEAMMEA
GWDPCLQAPFHLTPTLYRTLNRNQGNQGAPAESREVLQDTVNLLFNHPRQRNASRENLNLEPE
QPATGQPRSRPLKVAGSPTGRLAGDQGSEEAPQRPPASSATLRRQRHLNGKVSPEKESGPRQ
ILRSLVRLSVAFAERNPVEELTVDSPPVQQISQLSLLHQGQFQPKPNHRGNKYLAKPGGS
RSAIPDTDGPSARAGGQTDPEQEEGPLDPEEDLSVKQLLEELSSLLDPSTGLALDRLSAPD
PAWMARLSSLPLTTNYRDNVISPDAAATEEPRTFQTFGKAEAPELSPTGTRLASTFVSEMSSL
LEMLLEQRSSMPVEAASEALRRLSVCGRTLSLDLATSAASGMKVQGDPGGKTGTEGKSRGSS
SSSRCL
```

Important features:

Signal peptide:

amino acids 1-13

Transmembrane domain:

amino acids 719-739

N-glycosylation site.

amino acids 415-418, 582-585, 659-662, 662-665 and 857-860

Cadherins extracellular repeated domain signature.

amino acids 123-133, 232-242, 340-350, 448-458 and 553-563

FIGURE 172

CGGACGCGTGGCGGACGCGTGGGGAGAGCCGCAGTCCGGCTGCAGCACCTGGAGAAGG
CAGACCGTGTGAGGGGGCCTGTGGCCCCAGCGTGTGGCCTGGGGAGTGGAAAGTGGAG
GCAGGAGCCTCCTTACACTCGCCATGAGTTCCTCATCGACTCCAGCATCATGATTACCT
CCCAGATACTATTTTGAGTTGGCTTGTACAGGTGATCTTCTCCGTGACGTTGCATTTCCTGCACC
GAGATACGTCAGTATGTTGTACAGGTGATCTTCTCCGTGACGTTGCATTTCCTGCACC
GTTTGAGCTCATCATTTGAAATCTTAGGAGTATTGAATAGCAGCTCCGTTATTTCACT
GGAAAATGAACCTGTGTGAATTCTGCTGATCCTGGTTTCATGGTGCCTTTACATTGGC
TATTTATTGTGAGCAATATCCGACTACTGCATAAACAAACGACTGCTTTTCCTGTCTT
ATGGCTGACCTTATGTATTCTGGAAACTAGGAGATCCCTTCCCATTCTCAGCCAA
AACATGGATCTTATCCATAGAACAGCTCATCAGCCGGTTGGTGTGATTGGAGTGACTCTC
ATGGCTCTTCTTCTGGATTGGTGTCAACTGCCATACACTTACATGTCTTACTCCT
CAGGAATGTGACTGACACGGATATTCTAGCCCTGGAACGGCGACTGCTGCAAACCATGGATA
TGATCATAAGCAAAAGAAAAGGATGGCAATGGCACGGAGAACATGTTCCAGAAGGGGAA
GTGCATAACAAACCATCAGGTTCTGGGAATGATAAAAGTGTACCACCTCAGCATCAGG
AAAGTAAAATCTTACTCTTATTCAACAGGAAGTGGATGCTTGGAAAGAATTAAGCAGGCAGC
TTTTCTGGAAACAGCTGATCTATGCTACCAAGGAGAGAACATGTTCCAGAAGGGGAA
AAGGGAAATATTTAATTTCTGGTTACTTTCTCTATTACTGTGTTGGAAAATTT
CATGGCTACCATCAATATTGTTTGATCGAGTTGGAAAACGGATCCTGTACAAGAGGCA
TTGAGATCACTGTGAATTATCTGGGAATCCAATTGATGTGAAGTTGGTCCAACACATT
TCCTTCATTCTTGGAAATAATCGTCACATCCATCAGAGGATTGCTGATCACTTTAC
CAAGTTCTTATGCCATCTCTAGCAGTAAGTCCTCCAATGTCAATTGTCCTGCTATTAGCAC
AGATAATGGCATGTACTTTGTCTCCTCTGTGCTGATCCGAATGAGTATGCCTTAGAA
TACCGCACCATAATCACTGAAGTCCTGGAGAACTGCAGTTCAACTCTATCACCGTTGGTT
TGATGTGATCTCCTGGTCAGCGCTCTCTAGCATACTCTCCTCTATTGGTCACAAAC
AGGCACCAGAGAACGAAATGGCACTTGAACTTAAGCCTACTACAGACTGTTAGAGGCCAGT
GGTTCAAAATTAGATATAAGAGGGGGAAAATGGAACCAGGGCCTGACATTATAAAC
AAACAAAATGCTATGGTAGCATTTCACCTCATAGCATACTCCTCCCCGTAGGTGATA
CTATGACCATGAGTAGCATGCCAGAACATGAGAGGGAGAACTAACTCAAGACAATACTCA
GCAGAGAGCATCCCGTGTGGATATGAGGCTGGTAGAGGCGGAGAGGAGCCAAGAAACTAA
AGGTGAAAATACACTGGAACTCTGGGCAAGACATGTCTATGGTAGCTGAGCCAACACGT
AGGATTCCGTTAAGGTTACATGGAAAAGGTTAGCTTGCCTGAGATTGACTCATT
AAAATCAGAGACTGTAACAAAAAAAAAAAAAGGGCGGCCGACTCTAGAGTCG
ACCTGCAGAAGCTTGGCCGCCATGGCCCAACTGTTATTGCAGCTTATAATG

FIGURE 173

MSFLIDSSIMITSQILFFGFGWLFFMRQLFKDYEIRQYVVQVIFSVTFAFSCTMFELIIFEI
LGVLNSSSRYFHWKMNLCVILLILVFMVPFYIGYFIVSNIRLLHKQRLLFSCLLWLTMYFF
WKLGDPPILSPKHGILSIEQLISRGVIGVTLMALLSGFGAVNCPTYMSYFLRNVTDTDI
LALERRLLQTMDMIISKKRMMARRTMFQKGEVHNKPSGFWMKSVTSASGSENLTLIQ
QEVDALLELSRQLFLETADLYATKERIEYSKTFKGKYFNFLGYFFSIYCVWKIFMATINIVF
DRVGKTDPVTRGIEITVNYLGIQFDVKFWSQHISFILVGIIIVTSIRGLLITLTKFFYAISS
SKSSNVIVVLLAQMIMGMYFVSSVLLIRMSMLEYRTIITEVLGELQFNFYHRWFDVIFLVSA
LSSILFLYLAHKQAPEKQMAP

Important features:

Signal peptide:

amino acids 1-23

Potential transmembrane domains:

amino acids 37-55, 81-102, 150-168, 288-311, 338-356, 375-398,
425-444

N-glycosylation sites.

amino acids 67-70, 180-183 and 243-246

Eukaryotic cobalamin-binding proteins

amino acids 151-160

FIGURE 174

CATGGGAAGTGGAGCCGGAGCCTCCTTACACTGCCATGAGTTCCATCGACTCCAGCA
TCATGATTACCTCCCNGANACTATTTTGAGTTGGCTTCTTCCNGCGCAATGTT
TAAAGACTATGAGATACTGTCAGTATGTTGACNGGTGATCTCTCCGTGACGTTGCCATT
CTTGCACCAGTTGAGCTCATCTTGAAATCTNGGAGTATTGAATAGCAGCTCCGT
TATTTCACTGGAAAATGAACCTGTGTGTAATTCTGCTGATCCTGGTTNTCATGGTGCCTT
TTACATTGGCTATTTATTGTGAGCAATATCCGACTACTGCATAAACAAACGACTGCTTTT
CCTGTCTCTTATGGCTGACCTTATGTATTCCAG

FIGURE 175

GTGTTGCCCTGGGGAGGGGAAGGGGAGCCNGCCCTTCCTAAAATTGGCCAAGGGTTTC
TTTNTTGAATTCCGGTTNNGNATACCTTCCCAGAAAATATTTTGAGTTGGTAGNTT
TTTTCATGCGCCAATTGTTAAAGACTATGAGATACGTCACTATGTTGTACAGGTGATNTT
NTCCGTGACGTTGCATTTCTGCACCATGTTGAGCTCATCATNTTGAAATNTTAGGAG
TATTGAATAGCAGCTCCGTTATTCACGGAAAATGAACCTGTGTGAATTCTGCTGATC
CTGGTTTCATGGTGCCTTTTACATTGGCTATTTATTGTGAGCAATATCCGACTACTGCA
TAAACAAACGACTGCTTTCTGTCTNTTATGGCTGACCTTATGTATTNTNTGGAAAN
TAGGAGATCCCTTCCCATTCTC

FIGURE 176

CTCGCGCAGGGATCGTCCC**ATGGCCGGGGCTCGGAGCCCGCACCC**TTGGGGGGCTCCGGGATTGCTACCTTT
TGGCTCCCTGCTCGTCGA**ACTGCTTCTCACGGCGTCGCC**TTCAATCTGAGCTGATGGTGCTCGC
GGAGGGCGAGCCAGGCA**GGCTCTT**CGGCTCTCTGTCAGGGCAGTGCAGGCCAGGGAGCTGAA
GCTGCTGGTGGGTGCT**CCCAGGGCTGGCTCTCT**GGGGCAGCAGGCAATCGCACTGGAGGCTCTCGCTTG
CCC**GGTGA**GGCTGGAGGAGACTGACTGCTACAGAGTGGACATCGACCAGGGAGCTGATATGCAA
GGAGAACCAGTGGTGGGAGTCAGTGTGGAGCCAGGGGGCTGGGGCAAGATTGTA**CTGTG**ACACCGATA
TGAGGCAAGGCAGCAGTGGGACAGATCTGGAGACGGGGATATGATTGTCGCTGTTGTGCTCAGCCAGGA
CCTGGCCAT**CCGGGATGAGTGGATGGG**ATGGAAGTTCTGTGAGGGACGCCCAAGGCCATGAACAA
TGGGTTCTGCCAGCAGGGCACAGCTGCCCTCTCCCTGATAGCCACTACCTCCTTTGGGCCCCAGGAAC
CTATAATTGGAAGGGCACGGCCAGGGTGGAGCTGTGACAGGGCTCAGCGGACCTGGCACACCTGGACAGG
TCCTACAGGGCGGGGGAGAGAAGGAGCAGGACCCCGCCTCATCCGGTCCGCCAACAGCTACTTGGCTT
CTCTATTGACTCGGGAAAGGTCTGGTGCCTGAGAAGAGCTGAGCTTGTGGCTGGAGCCCCCGGCCAACCA
CAAGGGTGTGTTGCTACCTGCTGAGCTGACCTCAACAGTGTATGGCTGGCCAGACCTGATA**GTGGG**
TCCCCCTACTTCTTGAGCGCAAGAAGAGCTGGGGGTGCTGTATGTACTTGAACCAGGGGTCACTG
GGCTGGGATCTCCCCTCTCCGGCTCGGGCTCCCTGACTCCATGTTGGGATCAGCCTGGCTGTCCGGG
CCTCAACCAAGATGGCTTCCAGATATTGAGTGGGTCGCCAACCTTACAGGTGCTGGAGGGCAGGTTATGCT
TGGGAGCAGCTGGGGTTGCTGCCAACCTTACAGGTGCTGGAGGGCAGGCTGGGATCAAGAGCTTCGG
CTACTCCCTGTCAAGCAGCTTGGATATGGATGGAACCAATACCTGACCTGCTGGTGGCTCCCTGGCTGACAC
CGCAGTGCTTCAAGGGCAGACCCATCCTCATGTCCTCCATGAGGTCTATTGCTCCACGAAGCATCGACCT
GGAGCAGCCA**ACTGTGCTGGGGCA**CTCGGTCTGTGGACACTAAGGGTCTGTTCA**GT**CTACATTGCA
CAGCAGCTATAGCC**ACTGTGGCC**CTGGACTATGTGTTAGATGCGGACACAGCAGGGAGGCTCGGGGG
CTCCCGTGTGACGTTCTGTGAGCGTAA**CTGGAAAGAACCC**AAGCACCAGGGCTCGGGCACCGTGTGG
CCAGCATGACCGAGCTGTGGAGACGCCATGTTCCAGCTCAGGAAAGCTTCAAGAACAGCTTGGGG
AGTGA**CTTGT**CTACAGTCTCCAGACCCCTCGGCTCGGCAGAGGCTCTGGCCAGGGCTGCCTCCAGTGG
CCCCATCCTCAATGCCAACAGGGCAGACCCAGGGCAGAGAGATCCACTCTGTGAA**AGCAAGG**CTGTGG
CAAGATCTGCCAGAGCAATCTGAGCTGGTCCACGCCGCTTCTGTACCCGGTCA**CG**GA**AC**CGGAATT
CTGCCCATGGATGTGGATGGAACAAACAGCCCTGTTGACTGAGTGGGAGGCCAGTCATTGGCCTGGAGCTGAT
GGTACCAACCTGCCATGGACCCAGCCAGGCCAGGGCTGATGGGATGATGCCATGAAGGCCAGCTCTGGT
CATGCTTCTGACTCACTGCA**CTACTCAGGGTCCGGCC**CTGGACCCCTGCCAGAGCAAGCCACT
TGAGAATGCTCCCATGTTGAGTGTGAGCTGGGAACCCATGAAGAGAGGTGCCAGGTACCTTCTACCTCAT
CCTAGCACCTCCGGATCAGCATTGAGACCACGGAACTGGAGGGTAGAGCTGCTGTTGCCACGATCAGTGAGCA
GGAGCTGCATCCAGTCTCGCACGAGCCGTCTTATTGAGCTGCCACTGTCATTGCCAGGAATGCCATTCC
CCAGCAACTCTTCTCTGGTGTGGAGGGCGAGAGAGGCCATGAGCTGTGAGCGGGATGTGGCAGCAAGGT
CAAGTATGAGGTACGGTTCCAACCAAGGCCAGTCCTCAGAACCCCTGGCTCTGCCCTCAACATCATGTG
GCCATGAGATTGCCATGGGAAGTGGTGTGACCCATGAGGGTAGAGCTGGAGGGCGGGCAGGGCCTGG
GCAGAAAGGGCTTGTCTCCAGGCCAACATCCTCACCTGGATGTGGACAGTAGGGATAGGAGGCGGG
GCTGGAGCCACCTGAGCAGCAGGAGCCTGGTAGCGGCAGGAGGCCAGCATGTCCTGGCCAGTGTCTCTGC
TGAGAAGAAGAAA**ACATCACCC**TGGACTGCGCCGGGCACGCCA**ACTGTG**GTGGTGTTCAGCTGCC
CAGCTTGTACCGCGCGGGCTGTGTCGATGTCCTGGGGCGTCTGGAAACAGCAGCCTCTGGAGGAGTACTCAGC
TGTGAAGTCCTGGAAAGTGA**TTGT**CCGGGCAACATCACAGTGAAGTCTCCATAAAAGAAC**TTG**ATGTCCTGG
TGCCTCCACAGTGA**CTCCAGTGT**GGTATACTGGACCCCATGGCTGGCAGAGGAGTGCCCTGGGG
CATCCTCTGGCTGTACTGGCTGGCTGCTGGTGTAGCACTGTCGGTGTCTGGAAAGATGGATTCTT
CAAACGGGCAAGCACCCAGGGCAGGGCAGCGTCCCTGAGGAA**ACTGGG**CAGCCGGGGAGGGCCGG
GTTCAGGGAGAGAACAGGGCAGCATCCTGAGGAACAA**ACTGGG**CAGCCGGGGAGGGCCGG
CCCCATCTGGCTGCTGACGGCAGCTGGGCCCGATGGGATCAGGCCAGGCC**TAG**TTCC
CATGCTCCAGCCTGGCTGTGGCTGCCCTCCATCCCTCCAGAGATGGCTCTGGGATGAAGAGGGTAGAGT
GGGCTGCTGGTGTGCAAGATTGGCAGGATGGCTTCTCAGGGGACAGACCTCTCCACCCACAAGAAC
TCCCTCCACCA**ACTTCCC**TTAGAGTGTGAGATGAGAGTGGTAAATCAGGGACAGGGCATGGGTAGGG
TGAGAAGGGCAGGGGTGCTGTGATGCAAAGGTGGGAGAAGGGATCTA**ACCTAGAGGG**TGCGGGAGGG
GTAACAGGACCCCAAGGAC**CTGCC**CTCCCCGAAGTGCCTTAACCTAGAGGGTGCAGGGAGGG
CTCAGGCTGCTCTCTAGTTCCCCTCATCTGACCTTAGTTGCTGCCATCAGTCTAGTGGTTCTGTGG
TTCGTCTATTATTAAAAA**ATTTGAGAAC**AAAAAAAAAAAAAA

FIGURE 177

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA55737

><subunit 1 of 1, 1141 aa, 1 stop

><MW: 124671, pI: 5.82, NX(S/T): 5

MAGARSRDPWGASGICYLFGSLLVELLFSRAVAFNLDVMGALRKEGEPGSLFGFSVALHRQL
QPRPQSWLLVGAPQALALPGQQANRTGGLFACPLSLEETDCYRVDIDQGADMQKESKENQWL
GVSVRSGPGGKIVTCAHRYEARQRVDQILETRDMIGRCFVLSQDLAIRDELDGEWFCKCEG
RPQGHEQFGFCQQGTAAAFSPDSHYLLFGAPGTYNWKGTAIRVELCAQGSADLAHLDGPYEA
GGEKEQDPRLIPVPMNSYFGFSIDSGKGLVRAEELSFVAGAPRANHKGAVVILRKDSASRLV
PEVMLSGERLTSGFGYSLAVADLNSDGWPDLIVGAPYFFERQEELGGAVYVYLNQGGHWAGI
SPLRLCGSPDSMFGISLAVLGDLNQDGFPDIAVGAPFDGDGKVFYHGSSLGVVAKPSQVLE
GEAVGIKSFYSLSGSLDMDGQNQYPDLLVGSЛАDTAVLFRARPILHVSHEVSIAPRSIDLEQ
PNCAGGHSVCVDLRVCFSYIAVPSSYSPVALDYVLDADTDRLRGQVPRVTFLSRNLEEPK
HQASGTVWLKHQHDRVCGDAMFQLQENVKDKLRAIVVTLQSYSLQTPRLRRQAPGQGLPPVAP
ILNAHQPSTQRAEIHFLKQGCGEDKICQSNLQLVHARFCTRSDTEFQPLPMDVDGTTALFA
LSGQPVIGLELMVTNLPSPDAQPQADGDDAHEAQLLVMLPDSLHYSGVRALDPAEKPLCLSN
ENASHVECELGPNPMKRGAAQVTFLILSTSGISIETTELEVELLLATISEQELHPVSARARVF
IELPLSIAGMAIPQQLFFSGVVRGERAMQSERDVGSKVKYEVTVSNQGQLRTLGS AFLNIM
WPHEIANGKWLLYPMQVELEGGQGPGQKGLCSPRPNILHLDVDSRDRRRRELEPPEQQEPGE
RQEPMSSWPVSSAEKKKNITLDCARGTANCVFSCPLYSFDRAAVLHWGRLWNSTFLEEY
SAVKSLEVIVRANITVKSSIKNLMRDASTVIPVMVYLDPMMAVVAEGVPWWVILLAVLAGLL
VLALLVLLWKMGFFKRAKHPEATVPQYHAVKIPREDRQQFKEEKTGTILRNNWGSPRREGP
DAHPILAADGHPELGPDPGHPGPGTA

Important features:

Signal peptide:

amino acids 1-33

Transmembrane domain:

amino acids 1040-1062

N-glycosylation sites.

amino acids 86-89, 746-749, 949-952, 985-988 and 1005-1008

Integrins alpha chain proteins.

amino acids 1064-1071, 384-408, 1041-1071, 317-346, 443-465, 385-407, 215-224, 634-647, 85-99, 322-346, 470-479, 442-466, 379-408 and 1031-1047

FIGURE 178

CGCGCCGGCGCAGGGAGCTGAGTGGACGGCTCGAGACGGCGCGCGTGCAGCAGCTCCAGA
AAGCAGCGAGTTGGCAGAGCAGGGCTGCATTCCAGCAGGAGCTGCAGCACAGTGCTGGCT
CACAAACAAGATGCTCAAGGTGTCAAGCTCAGCCGTACTGTGTGTGCAGCCGCTTGGTGCAGTCA
GTCTCTCGCAGCTGCCGCGCGTGGCTGCAGCCGGGGGGCGGTGGACGGCGGTAACTTC
TGGATGATAAACAAATGGCTACCAACAATCTCTCAGTATGACAAGGAAGTCGGACAGTGGAAC
AAATTCCGAGACGAAGTAGAGGATGATTATTCCGCACTTGGAGTCCAGGAAAACCCTCGA
TCAGGCTTAGATCCAGCTAAGGATCCATGCTTAAAGATGAAATGTAGTCGCCATAAGTAT
GCATTGCTCAAGATTCTCAGACTGCAGTCTGCATTAGTCACCGGAGGCTTACACACAGGATG
AAAGAACGAGGAGTAGACCATAGGCAGTGGAGGGTCCCATTATCACCTGCAAGCAGTG
CCCAGTGGCTATCCCAGCCCTGTTGTGGTTCAGATGGTCATACTACTCTTCAGTGCA
AACTAGAATATCAGGCATGTGCTTAGGAAAACAGATCTCAGTCAAATGTGAAGGACATTGC
CCATGTCCTTCAGATAAGCCCACCAAGTACAAGCAGAAATGTTAAGAGAGCATGCAGTGACCT
GGAGTTCAAGGAAGTGGCAAACAGATTGCGGGACTGGTCAAGGCCCTCATGAAAGTGGAA
GTCAAAACAAGAACAGACAAAACATTGCTGAGGCCTGAGAGAACAGATTGATACACCAGCATC
TTGCCAATTGCAAGGACTCACTGGCTGGATGTTAACAGACTTGATACAAACTATGACCT
GCTATTGGACCAGTCAGAGCTCAGAACAGCTTACCTTGATAAGAACAGTGTACCAAGG
CATTCTCAATTCTGTGACACATACAAGGACAGTTAACATCTAATAATGAGTGGTGTAC
TGCTTCCAGAGACAGCAAGAACCCACCTGCCAGACTGAGCTCAGCAATATTCAAGCAGGGCA
AGGGTAAAGAACGCTCTAGGACAGTATATCCCCCTGTGTGATGAAGATGGTTACTACAAGC
CAACACAATGTCATGGCAGTGGACAGTGCTGGTGTGACAGATATGAAATGAAGTC
ATGGGATCCAGAATAAATGGTGTGCAAGATTGCTATAGATTGAGATCTCCGGAGATTT
TGCTAGTGGCGATTTCATGAATGGACTGATGATGAGGATGATGAAGACGATATTGAATG
ATGAAGATGAAATTGAAGATGATGATGAAGATGAAGGGATGATGATGATGGTGGTGTGAC
CATGATGTATACATTTGATTGATGACAGTTGAAATCAATAATTCTACATTCTAATATT
CAAAATGATAGCCTATTAAAATTCTTCTCCCAATAACAAATGATTCTAAACCTCA
CATATATTGTATAATTATTGAAAAATTGCAGCTAAAGTTAGAACATTGTTAAAT
AAGAACATTTGCTTGAGTTTATATTCCCTACACAAAAAGAAAATACATATGCAGTCTA
GTCAGACAAAATAAGTTGAAGTGCTACTATAATAAATTTCACGAGAACAAACTTGT
AAATCTTCCATAAGCAAAATGACAGCTAGTGCTGGGATCGTACATGTTAATTGAAAG
ATAATTCTAAGTGAATTTAAAATAAAATTGACCTGGGTCTTAAGGATTAGG
AAAAATATGCATGCTTAATTGCATTTCAAAGTAGCATCTGCTAGACCTAGATGAGTCAG
GATAACAGAGAGATACCACATGACTCCAAAAA

FIGURE 179

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49829
><subunit 1 of 1, 436 aa, 1 stop
><MW: 49429, pI: 4.80, NX(S/T): 0
MLKVSAVLCVCAAWCSQSLAAAAAVAAAGGRSDGGNFLDDKQWLTTISQYDKEVGQWNKFR
DEVEDDYFRTWSPGKPFQALDPAKDPCLMKCSRHKVCIAQDSQTAVCISHRRLTHRMKEA
GVDHRQWRGPILSTCKQCPVVYPSPVCGSDGHTYSFQCKLEYQACVLGKQISVKCEGHCP
SDKPTSTS RNVKRACSDLEFREVANRLRDWFKALHESGSQNKKTKLLRPERSRFDTSILPI
CKDSLGMFNRLDTNYDLLLQSELRSIYLDKNEQCTKAFFNSCDTYKDSLISNNEWCYCFQ
RQDPPCQTELSNIQKRQGVKKLLGQYIPLCDEDGYYKPTQCHGSVGQCWCVDRYGNEVMGS
RINGVADCAIDFEISGDFASGDFHEWTDDDEDDIMNDEDEIEDDDEDEGDDDDGGDDHDVYI
```

Important features:

Signal peptide:

amino acids 1-16

Leucine zipper pattern.

amino acids 246-267

N-myristoylation sites.

amino acids 357-362, 371-376 and 376-381

Thyroglobulin type-1 repeat proteins

amino acids 353-365 and 339-352

FIGURE 180

CAGACTCCAGATTCCCTGTCAACCACGAGGGAGTCCAGAGAGGAAACGC GGAGCGAGACAACAGTACCTGACGC
CTCTTCAGCCGGATCGCCCCAGCAGGG**A**TGGCGACAAGATCTGGCTGCCCTCCCGTGCCTCTGGCC
GCTCTGCCTCCGGTGTGCTGCCCTGGGGCGCCGGCTCACACCTCCCTCGATAGCGACTTCACCTTACCCCT
CCCGCCGGCCAGAAGGAGTGTCTTACCCAGCCCAGTGGCCCTGCTGGAGATCGAGTACCAAGTTTA
GATGGAGCAGGATTAGATATTGATTCCATCTTGCCTCTCAGAAGGAAACCTTAGTTTGACAACAGAAGAAAA
TCAGATGGAGTTCACACTGTAGAGACTGAAGTTGGTATTACATGTTCTGCTTGACAATACATTAGCACCATT
TCTGAGAAGGTGATTTCTTGAAATTAACTGGATAATATGGAGAACAGGCACAAGAACAGAACAGATTGGAG
AAATATATTACTGGCACAGATATATTGGATATGAAACTGGAGAACATCCTGGAAATCCATCAACAGCATCAAGTCC
AGACTAACGAAAGTGGCACATACAAATTCTGCTTAGAGCATTGAAAGCTGTGATGAAACATACAAGAAC
AACTTTGATAGAGTCATTTCTGGTCTATGGTAAATTAGTGGTCAATGGTGGTGTGAGCCATTCAAGTTAT
ATGCTGAAGAGTCGTGTTGAAGATAAGAGGAAAGTAGAACTAAACTCCAAACTAGAGTACGTAACATTGAAA
AATGAGGCATAAAATGCAATAACTGTTACAGTCAAGACCATTAATGGCTCTCCAAAATATTGAGATATA
AAAGTAGGAAACAGGTATAATTAAATGTGAAAATTAAAGTCTTCACCTTCTGTGCAAGTAATCTGCTGATCCAG
TTGTAAGTGTGAAACAGGAATATTGAGGTTAACTGAATGAAGCCATATTAAATACTGCAT
TTCTCTAACTTTGAAAAATTGCAAATGTCTTAGGTGATTAAATAATGAGTATTGGGCTAATTGCAACACC
AGTCTGTTTAACAGGTTCTATTACCCAGAACCTTTGTAAATGCGGAGTTACAAATTAAACTGTGGAAGTT
TCAGTTTAAGTTAAATCACCTGAGAATTACCTAATGATGGATTGAATAATCTTAGACTACAAAGCCAA
CTTTCTCTATTACATATGCATCTCCTATAATGAAATAGATAATAGCTTGAATACATTAGGTTTTG
AGATTTTATAACCAAATACATTCACTGTAACATATTAGCAGAACGATTAGTCTTGACTTTGCTTACATT
CCAAAGCTGACATTTCACGATTCTAAAAACACAAAGTTACACTAAATTAGGACATGTTCTCTTTG
AAATGAAGAATATAGTTAAAAGCTCCTCCATAGGGACACATTTCCTAACCCTTAACAAAGTGTAGGA
TTTAAAATTAATGTGAGGTTAAATAAGTTATTAAATAGTATCTGTCAGTTAAATCTGTCAACAGTTAA
TAATCATGTTATGTTAAATTAAACATGATTGCTGACTTGGATAATTACCTTACCGAGTTATGAAGGAAATA
TTGCTAAAATGATCTGGGCTTACCATAAAATAAATCTCTTCTGAGCTCTAAGAATTATCAGAAAACAGGAA
AGAATTAGAAAACCTGAGAAAACCTAATCCAAAATAAAATTCACTTAAGTAGAACATATAAAATATCTAGA
ATCTGACTGGCTCATGACATCCTACTCATAAACATAAACTAAAGGAGATGATTATTCAGTTAGCTGGAA
AAACTTGGCTGTAGGTTTATTCTACAAGAATTCTGGTTGAATTATTTGTAAGCAGGTACATT
AAATGTAAGCCCTACTGTAAGGTTAGCACTGGGTGACATATTAAATTAAATTAAACACTTT
TAAAATGGCTTCTGAACACTTATTGATGTTGAAGTAGGATTAGAACATAGACTCCAAGTTAAA
CACCTAAATGTGAATAACCCATATACACAAAGTTCTGCCATCTAGCTTTGAAAGTCTATGGGGCTTAC
TCAAGTACTAGTAATTAACTTCATCATGAATGAACATATAATTAAAGTTATGCCATTATAACGTTTT
GAACATTTGAGTTAGAAACAAACTAAAATTGGGTATAGAACCCCTAACAGGTTAGTAATGCTGAAATT
CTTGATGAGCAATAATGATAACCAAGAGAGTGATTCACTCATAGTAGTATAAAAGAGATACTTCCC
TCTTAGGCCCTGGGAGAAGAGCAGCTTAGATTCCCTACTGGCAAGGTTAAAATGAGGAAATGCCGTAT
ATGATCAATTACCTTAATTGCCAAGAAAATGCTTCAGGTGTCTAGGGTATCCTCTGCAACACTGCAAGAACAA
AGGTCAATAAGATCCTGCCTATGAATACCCCTCCCTTGCCTGTTAAATTGCAATGAGAACAAATTACA
GTACCATAACTAATAAGCAGGGTACAGATATAACTACTGCATCTTCTATAAAACTGTGATTAAGAATTCTA
CCTCTCCTGTATGGCTGTTACTGTACTCTGACTCTTACCTAACATGAATTGTTACATAATTCTCT
ACATGTATGATTGTCCTGCACTGATCTAAACCTATGATTCACTTAACCTTCTACCATATAAAACGATAATTGCTT
TATTGAAAAGAATTAGGAATACTAAGGACAATTATTTATAGACAACAAACTAAAAGACAGATATTAAAGAGG
CATACCAAAAAAGCAAACCTGTAACAGAGTAAAGTAAATCTTAAATTTCTAAAGACATACTGTTATCTGCTT
CATATGCTTTTTAATTCACTATTCCATTCTAAATTAAAGTTATGCTAAATTGAGTAAGCTGTTATCCT
AACAGCTATTGCTTTCAATATAACAAATTAAAGTTAAACTACAAATTTAAACTACAAATTAACAGGCCAACGATTTC
CATATGTAGCAGTTACCGTGTTCACCTCACACTAACGGCCTAGAGTTGCTGTGATATGCAATTGGATGATTAAT
GTATGCTGTTCTTCATGTGAATGTCAAGACATGGAGGGTGTGTAATTGTTATGGTAAATTAAATCCTCTTA
CACATAATGGTGTCTAAAATTGACAAAAAAATGAGCACTTACAATTGATGTTCTCCTCAAATGAAGATTCTT
GTGAAATTGTTAAAGACATTGATTCCGCATGTAAGGATTTCATCTGAAAGTACAATAATGCAACATCAGTGTG
CTCAAAACTGTTTATACCTATAAACAGCCATCTTAAATAAGCAACGTATTGTGAGTACTGATATGTATATAATAA
AAATTATCAAAGGAAA

FIGURE 181

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52196
><subunit 1 of 1, 229 aa, 1 stop
><MW: 26017, pI: 4.73, NX(S/T): 0
MGDKIWLPPVLLAALPPVLLPGAAAGFTPSLDSDFTFTLPAGQKECFYQPMPPLKASLEIEY
QVLDGAGLDIDFHLASPEGKTLVFEQRKSDGVHTVETEVGDYMFCDNTFSTISEKVIFFEL
ILDNMGEQAQEQQEDWKKYITGTDILDLMKLEDILESINSIKSRLSKSGHIQILLRAFEARDRN
IQESNFDRVNFWSMVNLVVVVSAIQVYMLKSLFEDKRKSRT
```

Important features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 195-217

N-myristoylation site.

amino acids 43-48

Tyrosine kinase phosphorylation site.

amino acids 55-62

FIGURE 182

CCATCCCTGAGATCTTTATAAAAAACCCAGTCTTGCTGACCAGACAAAGCATACCAGAT
CTCACCGAGAGTCGAGACACTATGCTGCCTCCATGCCCTGCCAGTGTGTCCTGGATG
CTGCTTCTGCCTCATTCTCCTGTCAAGGTTCAAGGTGAAGAAACCCAGAAGGAAGTGC
CTCTCCACGGATCAGCTGCCAAAGGCTCAAGGCCTATGGCTCCCCCTGCTATGCCTTGT
TTTGTACCAAAATCTGGATGGATGCAGATCTGGCTGCCAGAAGCGGCCCTGGAAAAA
CTGGTGTCTGTGCTCAGTGGGCTGAGGGATCCTCGTGTGTCCTCCCTGGTGAGGAGCATTAG
TAACAGCTACTCATACATCTGGATTGGCTCCATGACCCCACACAGGGCTCTGAGCCTGATG
GAGATGGATGGAGTGGAGTAGCACTGATGTGATGAATTACTTGCATGGAGAAAAATCCC
TCCACCCTTAAACCCCTGCCACTGTGGAGCCTGTCAAGAAGCACAGGATTCTGAAGTG
GAAAGATTATAACTGTGATGCAAAGTTACCCATGTCTGCAAGTCAAGGACTAGGGCAGGT
GGGAAGTCAGCAGCCTCAGCTGGCGTGCAGCTCATGGACATGAGACCAGTGTGAAGAC
TCACCCCTGGAAGAGAATATTCTCCCCAAACTGCCCTACCTGACTACCTGTCATGATCCTCC
TTCTTTCTTCTTCACCTTCATTCAGGCTTCTGTCTCCATGTCTTGAGATC
TCAGAGAATAATAATAAAAATGTTACTTATAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 183

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56965
<subunit 1 of 1, 175 aa, 1 stop
<MW: 19330, pI: 7.25, NX(S/T): 1
MLPPMALPSVSWMLLSCLILLCQVQGEETQKELPSPRISCPKGSKAYGSPCYALFLSPKSWM
DADLACQKRPSGKLVSVLSGAEGSFVSSLVRSISNSYSYIWIIGLHDPTQGSEPDGDGWEWSS
TDVMNYFAWEKNPSTILNPGHCGSLSRSTGFLWKDYNCDAKLPYVCKFKD
```

Important features:

Signal peptide:

amino acids 1-26

C-type lectin domain signature.

amino acids 146-171

FIGURE 184

CCAGTCTGTCGCCACCTCACTGGTGTCTGCTGTCCCCGCCAGGCAAGCCTGGGTGAGAGC
ACAGAGGAGTGGGCCGGGACCATGCGGGGGACGCCGGCTGGCGCTCCTGGCGCTGGTGTGGC
TGCCTGCGGAGAGCTGGCGCCGGCCCTGCGCTGCTACGTCTGTCCGGAGCCCACAGGAGTGT
CGGACTGTGTCACCACGCCACCTGCACCAACGAAACCATGTGCAAGACCACACTCTAC
TCCCAGGAGATAGTGTACCCCTCAGGGGACTCCACGGTACCAAGTCCTGTGCCAGCAA
GTGTAAGCCCTGGATGTGGATGGCATGGCCAGACCCCTGCCGTGTCCCTGCAATACTG
AGCTGTGCAATGTAGACGGGGCGCCGCTCTGAACAGCCTCCACTGCCGGCCCTCACGCTC
CTCCCACTCTTGAGCCTCCGACTGTAGAGTCCCCGCCACCCCCATGCCCTATGCCCA
GCCCGAATGCCTTGAAGAAGTGCCCTGCACCAGGAAAAAAAAAAAAAA

FIGURE 185

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56405
<subunit 1 of 1, 125 aa, 1 stop
<MW: 13115, pI: 5.90, NX(S/T) : 1
MRGTRLALLALVLAACGELAPALRCYVCPEPTGVSDCVTIATCTTNETMCKTTLYSREIVYP
FQGDSTVTKSCASKCKPSDVDGIGQTLPVSCCNTELNVGDGAPALNSLHCGALTLLPLLSLRL
```

Important features:

Signal peptide:

amino acids 1-17

N-glycosylation site.

amino acids 46-49

FIGURE 186

CTGCAGTCAGGACTCTGGGACCGCAGGGGGCTCCCGGACCCTGACTCTGCAGCCGAACCGGC
ACGGTTCTGGGGACCCAGGCTTGCAAAGTGACGGTCATTTCTCTTCTTCTCCCTCTT
GAGTCCTTCTGAGATGATGGCTCTGGCGCAGCGGGAGCTACCCGGGTCTTGTGCGATGG
TAGCGGCGGCTCTCGGCGGCCACCCCTCTGCTGGAGTGAGCGCCACCTTGAACTCGGTTCTC
AATTCCAACGCTATCAAGAACCTGCCCCCACCGCTGGCGGCTGCGGGCACCCAGGCTC
TGCAGTCAGCGCCGCCGGAAATCCTGTACCCGGGGAAATAAGTACCAGACCATTGACA
ACTACCAGCCGTACCCGTGCGCAGAGGACGAGGAGTGCGGCACTGATGAGTACTGCGCTAGT
CCCACCCCGGGAGGGACGCAGGCGTGCAAATCTGTCTCGCCTGCAGGAAGCGCCAAAACG
CTGCATGCGTCACGCTATGTGCTGCCCCGGGAATTACTGCAAAATGGAATATGTGTGCTT
CTGATCAAAATCATTCCGAGGAGAAATTGAGGAAACCATCACTGAAAGCTTGGTAATGAT
CATAGCACCTTGGATGGTATTCCAGAAGAACACCTTGTCTTCAAAATGTATCACACCAA
AGGACAAGAAGGTTCTGTTGTCTCCGGTCATCAGACTGTGCCTCAGGATTGTGTTGTGCTA
GACACTTCTGGTCCAAGATCTGTAAACCTGCTCTGAAAGAAGGTCAAGTGTGTACCAAGCAT
AGGAGAAAAGGCTCTCATGGACTAGAAATATTCCAGCGTTACTGTGGAGAAGGTCTGTC
TTGCCGGATACAGAAAGATCACCATCAAGCCAGTAATTCTTAGGCTTCACACTTGTCA
GACACTAAACCAGCTATCAAATGCAGTGAACCTCTTTATATAATAGATGCTATGAAAACC
TTTATGACCTTCATCAACTCAATCCTAAGGATATAAGTTCTGTGGTTCTAGTTCA
TCCAATAACACCTCCAAAACCTGGAGTGTAAAGAGCTTGTGTTCTTATGAAACTCCCTG
TGATTGCAGTAAATTACTGTATTGTAAATTCTCAGTGTGGCACTTACCTGTAAATGCA
AACTTTAATTATTTCTAAAGGTGCTGCACTGCCTATTTCTTGTGTTATGAAATT
TTGTACACATTGATTGTTATCTGACTGACAAATATTCTATATTGAACTGAAGTAAATCATT
TCAGCTTATAGTTCTAAAAGCATAACCCTTACCCATTAAATTCTAGAGTCTAGAACGCA
AGGATCTTGGAAATGACAAATGATAGGTACCTAAATGTAACATGAAAATACTAGCTATT
TTCTGAAATGTAATCTTAAATTATATTCCCTTAGGCTGTGATAGTTTGAA
AATAAAATTTAACATTTAAAAAAAAAA

FIGURE 187

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57530
<subunit 1 of 1, 266 aa, 1 stop
<MW: 28672, pi: 8.85, NX(S/T): 1
MMALGAAGATRVFVAMVAAALGGHPLLGVSATLNSVLNSNAIKNLPPPLGGAAGHPGSAVSA
APGILYPGGNKYQTIDNYQPYPKAEDEECGTDEYCASPTRGGDAGVQICLACRKRRKRCMRH
AMCCPGNYCKNGICVSSDQNHFRGEIEETITESFGNDHSTLDGYSRRTTLSSKMYHTKGQEG
SVCLRSSDCASGLCCARHFWSKICKPVLKEGQVCTKHRRKGSHGLEIFQRCYCCEGLSCRIQ
KDHHQASNSSLHTCQRH
```

Important features:

Signal peptide:

amino acids 1-23

N-glycosylation site.

amino acids 256-259

Fungal Zn(2)-Cys(6) binuclear cluster domain

amino acids 110-126

FIGURE 188

TGTGTTCCCTGCAGTCAGAATTGGGACNGCAGGGGTTCCCGGACCTGATTTCAGCGGA
ACGGGAAGGTTTGTGGGACCCAGGTTGAAATGACGGTCATTTTTTCTTCCTTCNG
GAGTCCTTNTGAGANGATGGTTGGCGCAGCGGGAGCTAACCCGGTTTGTNGCGATG
GTAGCGGCGGTTTCGGCGGCCACCTNTGCTGGGAGTGAGCGCCACCTGAATCGGTTTC
AATTCCAACGNTATCAAGAACCTGCCACCGNTGGCGGCCTGCGGGCACCCAGGNTT
TGCAGTCAGCGCCGCCGGAAATCCTGTACCCGGCGGAATAAGTACCAAGACCATTGACA
ATTACCAGCCGTACCCGTGCGCAGAGGACGAGGAGTGCGGCACTGATGAGTACTGCGCTAGT
CCCACCCGGAGGGANGCGGGCGTCAAATNTGNTNGCCTGCAGGAAGCGCCGAAAACG
CTGCATGCGTCANGCTATGTGCTGCCCGGGATTACTGCAAAATGGAATATGTGTNTT
CTGATCAAAATCATTCCGAGGAGAAATTGAGGAAACCATCACTGAAAGCTTGGTAATGAT
CATAGCACCTTGGATGGG

FIGURE 189

GAGGAACCTACCGGTACCGGCCGCGCGCTGGTAGTCGCCGGTGTGGCTCACCTCACCAATCCCGTGC CGCGCG
CTGGGCCGTCGGAGAGTGCCTGCTCTCCTGCACGCCGGTCTGGGCTCGGCCAGGCCGGTCCGCC
GGTTTGAGGATGGGGAGTAGCTACAGGAAGCGACCCCGCATGGCAAGGTATATTTGTGGAATGAAAAGGA
AGTATTAGAAATGAGCTGAAGACCATTCACAGATAATTTGGGACAGATTGTGATGCTGATTCAACCT
TGAAGTAATGTAGACAGAAGTCTCAAATTGCATATTACATCAACTGGAACCCAGCAGTGAATCTAATGTCAC
TTAAATCAGAACTTGCATAAGAAAGAGA**ATGGGAGTCTGGTAA**ATAAGATGACTATATCAGAGACTGAAAAG
GATCATTCTCTGTTCTGATAGTGTATGCCATTAGTGGGCACAGATCAGGATTTCACAGTTACTGG
AGTGTCAAACACTGCAAGCAGTAGAGAAATAAGACAAGCTTCAGAAATTGGCATTGAAGTTACATCCTGATAA
AAACCGAATAACCCAAATGCACATGGCATTAAAATAATAGAGCATATGAAGTACTCAAAGATGAAGA
TCTACGGAAAAGTATGACAATATGGAGAAAAGGGACTTGAGGATAATCAAGGTGCCAGTATGAAAGCTGGAA
CTATTATCCTTATGATTGGTATTATGATGATGATCCTGAAATCATAACATTGGAAAAGAGAATTGATGC
TGCTGTTAATTCTGGAGAACTGGGTTGTAAATTAACTCCCCAGGCTTACACTGCCATGATTAGCTCC
CACATGGAGAGACTTGCTAAAGAAGTGGATGGGTTACTCGAATTGGAGCTGTTACTGTGGTGTGATGATAA
GCTTGCCGAATGAAAGGAGTCAACAGCTATCCCAGTCTTCATTTCGGTCTGGAATGGCCCAGTGAATA
TCATGGAGACAGATCAAAGGAGAGTTAGTGAAGTTGCAATGCAGCATGTTAGAAGTACAGTGACAGAACTTG
GACAGGAAATTGTCACACTCCACAAACTGCTTTGCTGCTGGTATTGGCTGGCTGATCACTTTGTTCAA
AGGAGGAGATTGTTGACTTCACAGACACGACTCAGGCTTAGTGGCATGTTCTCAACTCATTGGATGCTAA
AGAAATATATTGGAAGTAATACATAATCTCAGATTGAAACTACTTCGGAAACACACTAGAGGATGTT
GGCTCATCATCGGTGGCTGTTATTGAAATTCAATGATCCTGAGCTGAAA
AAAAACTCTACTAAAATGATCATATTCAAGTTGGCAGGTTGACTGTTCTCTGCACCAAGACATCTGAGTAA
TCTGTATGTTTCAGCGTCTCTAGCAGTATTAAAGGACAAGGAACCAAAGAATATGAAATTCTCATCGAAA
GAAGATTCTATATGATATACTTGCTTTGCCAAAGAAAGTGTGAATTCTCATGTTACACGCTGGACCTCAA
TTTCCTGCCAATGACAAAGAACATGGCTTGTGATTCTTCTGCCCCCTGGTGTCCACCATGTCAGCTTACT
ACCAAGAGTTACGAAGAGCATCAAATCTCTTATGGCAGCTTAAGTTGGTACACTAGATTGTAAGTTCATG
GGACTCTGTAACATGTATAACATTCAAGCTTATCCAAACAGTGGTATTCAACAGTCCAAACATTCTAGAGTA
TGAAGGACATCACTGCTGAACAAATCTGGAGTTCATAGAGGATCTTATGAAATCCTCAGTGGTCTCCCTAC
ACCCACCACTTCAACGAACTAGTTACACAAAGAAAACAACAGAAGTCTGGATGGTGTGATTCTATTCTCGTG
GTGTCATCCTTGCCAAGTCTTAATGCCAGAATGGAAAAGAATGCCGGACATTAACGGACTGATCAACGTGG
CACTAGATTGCCAACAGTATCATTCTTGTGCCAGGAAAGGTTCAAAGATAACCTGAGATAAGATT
TCCCCCAAATCAAATAAGCTTACAGTATCACAGTTACAGGTTGGAATAGGGATGCTTATTCCCTGAGAA
CTGGGTCTAGGATTTCACCTCAAGTATCCACAGATCTAACACCTCAGACTTCAGTGGAAAGTTCTACAAGG
GAAAATCATTGGGTGATTGATTCTATGTCCTTGTGGACCTTGCAGAATTGCTCCAGAATTGAGCT
CTTGGCTAGGATGATTAAAGGAAAAGTGAAGAGCTGGAAAAGTAGACTGTCAGGCTTATGCTCAGACATGCCAGAA
AGCTGGGATCAGGGCTATCCAACACTGTTAAGTTTATTCTACGAAAGAGCAAAGAGAAAATTCAAGAAGAGCA
GATAAATACCAGAGATGCAAAGCAATCGCTGCCATTAAAGTGAAGAAACTCTCGAAATCAAGGCAA
GAGGAATAAGGATGAACT**TTG**ATAATGTTGAAGATGAAGAAAAGTTAAAAGAAATTCTGACAGATGACATCAG
AAGACACCTATTAGAATGTTACATTATGATGGGAATGAATGAACATTCTTAGACTTGCAGTTGACTGCCA
GAATTATCTACAGCACTGGTGTAAAAGAAGGGCTGCAAACCTTTCTGTAAGGGCCGGTTATAAAATATT
GACTTGCAGGCTATAATATGTTACACATGAGAACAGAATAGACTCATGTTATTCTGTTATTGCT
TTAACACCTTAAAGGATCTTAAACTTACAGGCCAGGAAACAGTGCAGCAGTATAATGTCACACAGTAAGTACAC
GACCATAGATTGCTGCCCCCTGACGGACTTATAATGTTCAGGTGGCTGGCTGACATGAGTCTGCTGTGCT
ATCTACATAATGCTAAGTTGTTAAAGTCCACCTTCCACGTTTTGGCTGACCTGAAAAGAGGTAAC
TAGTTTGGTCACTTGTCTCTAAAGCTATCCCTAACCATATATTATTTCTGTTTAAAGGACCCAT
GATGTGGCAGTAAACAAACCTGTTATGCTGTTATTGAGGAGATTCTCATGTTTCTTCTCTCA
AAGGTTGAAAAGCTTTAAATTCTCACAGGCCAGGAAACAGTGCAGCAGTATAATGTCACACAGTAAGTACAC
AAATTGAGCAACAGTAAGTGCACAAATTCTGTTAGAGTATCTAAATATGTTATCAAGTATTAGGTTCTATATT
TAGCAATTAACTGGGCATTGTTAGAGTATCTAAATATGTTATCAAGTATTAGGTTCTATATT
TGTGTTCATGTTATTCTGAAATTGCTTTCATAGAAATTCTCCACTGATAGTTGATTGAGGAGTCTAATAT
TTACATATTGCCCTCTGAACTTGTGTTGACCTGTTACCTTATTACATTGGGTTTCTTCATAGTTG
TTTCTACTCCTGTCAGTCTATTATTCAATAGGAAAATTACTTACAGGTTGTTTACTGTAGCTTAT
AATGATACTGTTAGTTATTCCAGTTACTAGTTACTGTCAGAGGGCTGCCTTTCTGAGATAAATATT
ACTGAAGTTATTGAAAGAAATCAAGTATATAAATCTAGGAAAGGGATCTCTAGTTCTGTTGTTAGA
CTCAAAGAATCACAAATTGTCAGTAACATGTTAGTTAGTTATAATTCAAGAGTGTACAGAATGGTAAA
CCAATCAGTCAGGAAAGAGGCTGCAACTTTCAAAAAA

FIGURE 190

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56439
<subunit 1 of 1, 747 aa, 1 stop
<MW: 86127, pI: 7.46, NX(S/T): 2
MGVWLNKDDYIRDLKRIILCFLIVYMAILVGTQDFYSLLGVSKTASSREIRQAFKKLALKL
HPDKNPNNPNAHDFLKINRAYEVLKDEDLRKKYDKYGEKGLEDNQGGQYESWNYYRYDFGI
YDDDPEIITLERREFDAAVNSGELWFVNFSYSPGCSHCHDLAPTRDFAKEVDGLLRIGAVNC
GDDRMLCRMKGVNNSYPSLFIFRSGMAPVKYHGDRSKESLVSFAMQHVRSTVTELWTGNFVNS
IQTAFAAAGIGWLITFCSKGGDCLTSQTRLRLSGMLFLNSLDAKEIYLEVIHNLPDFELLSAN
TLEDRLAHHRWLLFFHFGKNENSNDPELKKLKTLKKNDHIQVGRFDSSAPDICSNLVFQP
SLAVFKGQGTKEYEIHGKKILYDILAFAKESVNSHVTTLGPQNFANDKEPWLVDFFAPWC
PPCRALLPELRRASNLLYQQLKFGTLDCTVHEGLCNMYNIQAYPTTVFNQSNIHEYEGHHS
AEQILEFIEDLMNPSVSVSLPTTFNELVTQRKHNEWMVDFYSPWCHPCQVLMPEWKRMA
LTGLINVGSIDCQQYHSFCAQENVQRYPEIRFFPPKSNKAYQYHSYNGNRDAYSRLIWGLG
FLPQVSTDLPQTFSKVLQGKNHWVIDFYAPWCGPCQNFAPFELLARMIKGKVAGKVDC
QAYAQTCQKAGIRAYPTVKFYFYERAKRNFQEEQINTRDAKAIAAALISEKLETLRNQGKRNKDEL
```

Important features:

Endoplasmic reticulum targeting sequence.

amino acids 744-747

Cytochrome c family heme-binding site signature.

amino acids 158-163

Nt-dnaJ domain signature.

amino acids 77-96

N-glycosylation site.

amino acids 484-487

FIGURE 191

AGACAGTACCTCCTCCCTAGGACTACACAAGGACTGAACCAGAAGGAAGAGGACAGAGCAAA
GCCATGAACATCATCCTAGAAATCCTCTGCTTGTGATCACCATCATCTACTCCTACTTGGA
GTCGTTGGTGAAGTTTCATTCCCTAGAGGGAGAAAATCTGTGGCTGGGAGATTGTTCTCA
TTACTGGAGCTGGCATGGAATAGGCAGGCAGACTACTTATGAATTGCAAAACGACAGAGC
ATATTGGTTCTGTGGATATTATAAGCGCGGTGGAGGAACTGCAGCTGAGTGCCGAAA
ACTAGGCGTCAGTGCATGCGTATGTGGTAGACTGCAGCAACAGAGAAGAGATCTATCGCT
CTCTAAATCAGGTGAAGAAAGAAGTGGGTGATGTAACAATCGTGGTGAATAATGCTGGGACA
GTATATCCAGCCGATCTCTCAGCACCAAGGATGAAGAGATTACCAAGACATTGAGGTCAA
CATCCTAGGACATTGGATCACAAAAGCACTTCCATCGATGATGGAGAGAAATCATG
GCCACATCGTCACAGTGGCTCAGTGTGCAGGCCACGAAGGGATTCCCTACCTCATCCCATA
TGTCCAGCAAATTGCCGCTGTTGGCTTCACAGAGGTCTGACATCAGAACTTCAGGCCTT
GGGAAAAACTGGTATCAAAACCTCATGTCCTGCCAGTTTGTAATACTGGGTTCACCA
AAAATCCAAGCACAAGATTATGGCCTGTATTGGAGACAGATGAAGTCGTAAGAAGTCTGATA
GATGGAATACTTACCAATAAGAAAATGATTGGTCCATCGTATATCAATATCTTCTGAG
ACTACAGAAGTTCTCCTGAACGCGCCTCAGCGATTTAAATCGTATGCAGAATATTCAAT
TTGAAGCAGTGGTGGCCACAAAATCAAATGAAATGAATAAAATAAGCTCCAGCCAGAGATG
TATGCATGATAATGATATGAATAGTTCGAATCAATGCTGCAAAGCTTATTCACATTTT
TCAGTCCTGATAATATTAAAAACATTGGTTGGCACTAGCAGCAGTCACAGAACAGATTA
ATTACCTGTCTCCTGTTCTCAAGAATATTACGTAGTTTCATAGGTCTGTTTCCCT
TCATGCCTCTTAAAAACTTCTGTGCTTACATAACATACTAAAAGGTTTCTTAAGATAT
TTTATTTCCATTAAAGGTGGACAAAAGCTACCTCCCTAAAAGTAAATACAAAGAGAACT
TATTTACACAGGAAAGGTTAAGACTGTTCAAGTAGCATTCCAATCTGTAGCCATGCCACAG
AATATCAACAAAGAACACAGAACATGAGTGACAGCTAACAGAGATCAAGTTCAGCAGGCAGCTT
ATCTCAACCTGGACATATTAAAGATTGAGCTTACGATTTGAAAGATTCCCTAGCCTCTTCTGTAT
TCATTAGCCAAAACGGTGCAACTCTATTCTGGACTTTACTTGATTCTGTCTTCTGTAT
AACTCTGAAGTCCACCAAAAGTGGACCCCTATATTCCCTCCCTTTATAGTCTTATAAGA
TACATTATGAAAGGTGACCGACTCTATTAAATCTCAGAATTAAAGTTCTAGCCCCATGA
TAACCTTTCTTGTAAATTATGCTTCATATATCCTGGTCCCAGAGATGTTAGACAAT
TTTAGGCTAAAAATTAAAGCTAACACAGGAAAGGAACGTACTGGCTATTACATAAGAAA
CAATGGACCCAAGAGAAGAA

FIGURE 192

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56409
<subunit 1 of 1, 300 aa, 1 stop
<MW: 33655, pI: 9.31, NX(S/T): 1
MNIILEILLLITIIYSYLESLVKFFIPQRKSVAGEIVLITGAGHIGRQTTYEFAKRQSI
LVLWDINKRGVEETAAECRKLGVTAHAYVVDCSNREEIYRSLNQVKKEVGDVTIVVNNAGTV
YPADLLSTKDEEITKTFEVNILGHFWITKALLPSMMERNHGHIVTVASVCHEGIPYLI PYC
SSKFAAVGFHRGLTSELQALGKTGIKTSCLCPVFVNTGFTKNPSTRLWPVLETDEVVRSЛИD
GILTNKKMIFVPSYINIFLRLQKFLPERASAILNRMQNIQFEAVVGHKIKMK
```

Important features:

Signal peptide:

amino acids 1-19

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 30-33 and 58-61

Short-chain alcohol dehydrogenase family protein

amino acids 165-202, 37-49, 112-122 and 210-219

FIGURE 193

CGGCGGGGGCTGCGGGCGCGAGGTGAGGGGCGCGAGGTGAGGGGCGCGAGGTCCCCAGCAGG
ATGCCCGGCTCTGCAGGAAGCTGAAGTGAGAGGCCGGAGAGGGCCAGCCCGCCGGGC
AGGATGACCAAGGCCGGCTTCCGGCTGTCGGCTGGCTGGGTGGTCACTGATGACCT
GCTGATCATCGTGTACTGGGACAGCGCAGGCCGCGCACTCTACTTGACACGTCCTCT
CTAGGCCGCACACGGGCCGCGCTGCCACGCCCGGCCGACAGGGACAGGGAGCTCACG
GCCGACTCCGATGTCGACGAGTTCTGGACAAGTTCTAGTGCTGGCGTGAAGCAGAGCGA
CCTTCCAGAAAGGAGACGGAGCAGCCGCTGCCAGGCCCGGCCGACAGGGAGCGTGAAGAG
GCTACGACTGGTCCCCGCGCAGGCCCGCGCAGCCCAGACCAGGGCCGGCAGCAGGCCGAG
CGGAGGAGCGTGCTGCCGGCTCTGCCCAACTCCAGCCTGCCCTCCCCACCAAGGAGCG
CGCATTGACGACATCCCCAAGCTGGAGCTGAGCCACCTGATCGTGGACGACCGCACGGG
CCATCTACTGCTACGTGCCAAGGTGGCCTGCACCAACTGGAAGCGCGTGTGATGATCGTGTG
AGCGGAAGCCTGCTGCACCGCGGTGCGCCCTACCGCGACCCGCTGCGCATCCCGCGAGCA
CGTGCACAACGCCAGCGCACCTGACCTAACAAAGTTCTGGCGCCCTACGGGAAGCTCT
CCCGCCACCTCATGAAGGTCAAGCTCAAGAAGTACACCAAGTTCTCTCGTGCACGGG
TTCGTGCCCTGATCTCCGCCCTCGCAGCAAGTTGAGCTGGAGAACGAGGAGTTCTACCG
CAAGTTGCCGTGCCCATGCTGCCGTGTACGCCAACACACCAGCCTGCCCTGCC
GCGAGGCCCTCCCGCTGGCTCAAGGTGTCCTGCCAACCTCATCCAGTACCTGCTGGAC
CCGCACACGGAGAAGCTGGGCCCTCAACGAGCAGCTGGCGCAGGTGTACCGCCTCTGCC
CCCGTGCCAGATCGACTACGACTTCGTGGAGACTCTGGACGAGGACGCCGC
AGCTGCTGCAGCTACTCCAGGTGGACCGGAGCTCCGCTCCCCCGAGCTACCGAACAGG
ACCGCCAGCAGCTGGAGGAGACTGGTCCGAAGATCCCCCTGGCCTGGAGGCAGCAGCT
GTATAAAACTCTACGAGGCCGACTTGTCTCTCGCTACCCCAAGCCGAAACCTCCTCC
GAGACTGAAAGCTTCGCGTTGCTTTCTCGCGTGCCTGGAACCTGACGCACGCGCAGCTCC
AGTTTTTTATGACCTACGATTTGCAATCTGGCTTCACTCCACTGCCTCTATCC
ATTGAGTACTGTATCGATATTGTTTAAGATTAATATTTCAAGGTATTAAATACGA

FIGURE 194

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56112
<subunit 1 of 1, 414 aa, 1 stop
<MW: 48414, PI: 9.54, NX(S/T): 4
MTKARLFRLWLVLSVFMILLIIVYWDSAGAAHFYLHTFSRPTGPPPLPTPGPDRDRELTA
DSDVDEFLDKFLSAGVKQSDLPRKETEQPPAPGSMEESVRGYDWSPRDARRSPDQGRQQAER
RSVLRGFCANSSLAFPTKERAFDDIPNSELSHLIVDDRHGAIYCYVPKVACTNWKRVMIVLS
GSLLHRGAPYRDPLRIPREHVHNASAHLTFNKFWRRYGKLSRHLMKVKKYTKFLFVRDPF
VRLISAFRSKFELENEEFYRKFAVPMLRLYANHTSLPASAREAFRAGLKVSFANFIQYLLDP
HTEKLAPPNEHWQVYRLCHPCQIDYDFVGKLETLDDEAAQLLQQLQVDRQLRFPPSYRNRT
ASSWEEDWFAKIPLAWRQQLYKLYEADFVLFGYPK PENLLRD
```

Important features:

Signal peptide:

amino acids 1-31

N-glycosylation sites.

amino acids 134-137, 209-212, 280-283 and 370-373

TNFR/NGFR family cysteine-rich region protein

amino acids 329-332

FIGURE 195

TCGGGCCAGAATTGGCACGGAGGCGGCACGAGGGCGACGGCCTCACGGGCTTGAGGTGA
AAGAGGCCAGAGTAGAGAGAGAGAGACCGACGTACACGGGATGGCTACGGAACCGC
ATGCCGGGAAGGTGGTGGTCGTGACCGGGGCGGGCGGCATCGGAGCTGGATCGTGC
GCCTTCGTGAACAGCGGGGCCGAGTGGTTATCTCGGACAAGGATGAGTCTGGGGGCC
CCTGGAGCAGGAGCTCCCTGGAGCTGTCTTATCCTCTGTGATGTGACTCAGGAAGATGATG
TGAAGACCCCTGGTTCTGAGACCATCCGCCATTGGCCGCCTGGATTGTGTTGTAACAAAC
GCTGGCCACCACCCACCCCCACAGAGGCCTGAGGAGACCTCTGCCAGGATTCCGCCAGCT
GCTGGAGCTGAACCTACTGGGACGTACACCTTGACCAAGCTGCCCTCCCTACCTGCG
AGAGTCAGGGAAATGTCATCAACATCTCCAGCCTGGTGGGGCAATCGGCAGGCCAGGCA
GTTCCCTATGTGGCCACCAAGGGGCAGTAACAGCCATGACCAAAGCTTGGCCCTGGATGA
AAGTCCATATGGTGTCCGAGTCAACTGTATCTCCCAGGAAACATCTGGACCCCGCTGTGG
AGGAGCTGGCAGCCTTAATGCCAGACCCCTAGGGCCACAATCCGAGAGGGCATGCTGGCC
CCACTGGGCCGCATGGCCAGCCGCTGAGGTGGGGCTGCCAGTGTCCCTGGCCTCCGA
AGCCAACCTCTGCACGGCATTGAACGTGCTCGTGACGGGGGTGCAGAGCTGGGTACGG
GCAAGGCCAGTCGGAGCACCCCCGTGGACGCCCGATATCCCTCTGATTTCTCTCATT
CTACTTGGGGCCCCCTTCCTAGGACTCTCCACCCAAACTCCAACCTGTATCAGATGCAGC
CCCCAAGCCCTTAGACTCTAACGCCAGTTAGCAAGGTGCCGGTCACCCCTGCAGGTTCC
AAAAACGATTTGCAGCC

FIGURE 196

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56045
<subunit 1 of 1, 270 aa, 1 stop
<MW: 28317, pI: 6.00, NX(S/T): 1
MATGTRYAGKVVVVTGGGRGIGAGIVRAFVNNSGARVVICDKDESGGRALEQELPGAVFILCD
VTQEDDVKTLVSETIRRGRLDCVNNAGHHPPPQRPEETSAQGFRQLLENLNLGTYTLTKL
ALPYLRLKSQGNVINISSLVGAIGQAQAVPYVATKGAVTAMTKALALDESPYGVRVNCISPGN
IWTPPLWEELAALMPDPRATIREGMLAQPLGRMGQPAEVGAAAVFLASEANFCTGIELLVTGG
AELGYGCKASRSTPVDAPDIPS
```

Important features:

N-glycosylation site.

amino acids 138-141

Short-chain alcohol dehydrogenase family protein

amino acids 10-22, 81-91, 134-171 and 176-185

FIGURE 197

AGGCAGGAGCAGCTGCAGGCTGACCTGCAGCTGGCGGAATGGACTGGCCTCACAAACCTG
CTGTTCTTACCATTCATCTCCTGGGCTGGGCCAGCCCAGGAGCCCCAAAGCAA
GAGGAAGGGCAAGGGCGGCCTGGGCCCTGGCCCTGGCCCTCACCAGGTGCCACTGGACC
TGGTGTACGGATGAAACCGTATGCCGCATGGAGGAGTATGAGAGGAACATCGAGGAGATG
GTGGCCCAGCTGAGGAACAGCTCAGAGCTGGCCCAGAGAAAGTGTGAGGTCAACTGCAGCT
GTGGATGTCCAACAAGAGGAGCCTGTCTCCCTGGGCTACAGCATCAACCACGACCCAGCC
GTATCCCCGTGGACCTGCCGGAGGCACGGTGCCTGTCTGGGCTGTGAACCCCTTCACC
ATGCAGGAGGACCGCAGCATGGTGAGCGTGCCGGTGTTCAGCCAGGTTCTGTGCGCCGCCG
CCTCTGCCGCCACCGCCCCGCACAGGGCTTGCCGCCAGCGCGCAGTCATGGAGACCATCG
CTGTGGCTGCACCTGCATCTCTGAATCACCTGGCCCAGAAGCCAGGCCAGCAGCCGAGA
CCATCCTCTTGACCTTGTGCCAAGAAAGGCCTATGAAAAGTAAACACTGACTTTGAAA
GCAAG

0 1 2 3 4 5 6 7 8 9

FIGURE 198

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA59294
<subunit 1 of 1, 180 aa, 1 stop
<MW: 20437, pI: 9.58, NX(S/T): 1
MDWPHNLLFLLTISIFLGLGQPRSPKSKRKGQGRGPLAPGPHQVPLDLVSRMKPYARMEY
ERNIEEMVAQLRNSSELAQRKCEVNLQLWMSNKRSLSPWGYSINHDPSRIPVSDLPEARCLCL
GCVNPFTMQEDRSMSVPVFSQVPVRRRLCPPPRTGPCRQRAVMETIAVGCTCIF
```

Important features:

Signal peptide:

amino acids 1-20

N-glycosylation site.

amino acids 75-78

Homologous region to IL-17

amino acids 96-180.

FIGURE 199

GCGCCGCCAGGCGTAGGCAGGGTGGCCCTCGCTCCGCTTCCTGAAAAACCGGGGG
CGAGCGAGGCTGCGGCCGCTGCCCTCCCCACACTCCCGCCGAGAACGCTCGCTCG
GCGCCCAACATGGCGGGTGGCGCTCGGCCGAGCTAACGGCGCTCTGGCGCTGGAT
CGCGGCTGTGGCGCGACGGCAGGCCCGAGGAGGGCGCTGCCGCCGGAGCAGAGCCGG
TCCAGCCCCTGACCGCCTCAACTGGACGCTGGTATGGAGGGCGAGTGGATGCTGAAATT
TACGCCCATGGTGTCCATCCTGCCAGCAGACTGATTCAAAGATGGAGGGCTTGTCAAAGAA
TGGTGAATACTTCAGATCAGTGTGGGAAGGTAGATGTCATTCAAGAACCGAGGTTGAGTG
GCCGCTTCTTGTCAACCCTCCCAGCATTTTCAAGGATGGATATTCCGCCGT
TATCGTGGCCCAGGAATCTCGAACGACTCGAGAATTATATCTTAGAGAAGAAATGGCAATC
AGTCGAGCCTCTGACTGGCTGAAATCCCCAGCTCTCTAACGATGTCAGGAAATGGCTGGTC
TTTTAGCATCTCGCAAGATATGGCATCTCACAACTATTCACAGTGACTCTTGGAAATT
CCTGCTTGGTGTCTTATGTGTGGCTAGCCACCTGGTTTGGCCTTTATGGG
TCTGGTCTGGTGTGAATATCAGAATGTTCTATGTGCCACTTCAAGGCATTATCTGAGC
GTTCTGAGCAGAATCGGAGATCAGAGGAGGCTCATAGAGCTGAACAGTTGCAGGATGCCAG
GAGGAAAAAGATGATTCAAATGAAGAAGAAAACAAAGACAGCCTGTAGATGATGAAGAAGA
GAAAGAAGATCTGGCGATGAGGATGAAGCAGAGGAAGAGGAGGAGGACAACCTGGCTG
CTGGTGTGGATGAGGAGAGAAGTGAGGCCATGATCAGGGCCCCCAGGAGAGGACGGTGTG
ACCCGGGAGGAAGTAGAGCCTGAGGAGGCTGAAGAACGGCATCTCTGAGCAACCTGCCAGC
TGACACAGAGGTGGAGACTCCTTGAGGCAGCGTAAAGTCAGCATGCTGACAAGGGAC
TGAGATTAAATGATGCGTTTCAAGAATACACACCAAAACAATATGTCAGCTCCCTTG
CCTGCAGTTGTACCAAATCCTTAATTTCTGAATGAGCAAGCTCTCTTAAAGATGCT
CTCTAGTCATTGGTCTCATGGCAGTAAGCCTCATGTATACTAAGGAGAGTCTCCAGGTGT
GACAATCAGGATATAGAAAAACAAACGTAGTGTGGATCTGTTGGAGACTGGATGGGAA
CAAGTTCATTTACTTAGGGTCAGAGAGTCTGACCAGAGGAGGCCATTCCAGTCCTAAATC
AGCACCTCCAGAGACAAGGCTGCAGGCCCTGTGAAATGAAAGCCAAGCAGGAGCCTGGCT
CCTGAGCATCCCCAAAGTGTAAACGTAGAACGCTTGCACTCTTCTTGTAAAGTATTAT
TTTGTCAAATTGCAAGGAAACATCAGGCACACAGTCATGAAATCTTCACAGCTAGAA
ATTGAAAGGGCTGGGTATAGAGAGCAGCTCAGAAGTCATCCCAGCCCTGTAAATCTCCTG
TGCTATGTTTATTCTTACCTTAATTTCCAGCATTCCACCATGGCATTCAAGGTCT
CCACACTCTTCACTATTATCTCTGGTCAGAGGACTCCAATAACAGCCAGGTTACATGAAC
TGTGTTGTTCAATTGACCTAACGGGTTAGATAATCAGTAACATAACCCCTGAAGCTGT
GACTGCCAACATCTCAAATGAAATGTTGGCCATCAGAGACTCAAAGGAAGTAAGGATT
TTACAAGACAGATTAAAAAAATTGTTGTCAAATATAGTTGTTGATTTTT
AAGTTTCTAAGCAATATTTCAGGCCAGAAGTCCTCTAAGTCAGTGTGGCTTGTAAAGGTAGT
CTTGTGAAGAAAAGTTGAATACTGTTGTTGTTCAAGGGTCTGGCTTGTAAAC
TACTTTAAATAACTAAAAACCACTCTGATTTCCCTCAGTGTGCTTGTGAAAG
GAATTAAATGAACTCCAGTACCTGAAAGTGAAGATTGATTTGTTCCATCTCTGTAATC
TTCCAAAGAATTATATCTTGTAAATCTCTCAATAACTCAATCTACTGTAAGTACCCAGGGAG
GCTAATTCTT

FIGURE 200

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56433
<subunit 1 of 1, 349 aa, 1 stop
<MW: 38952, pI: 4.34, NX(S/T): 1
MAGGRCGPQLTALLAAWIAAVAAATAGPEEAALPPEQSRVQPMTASNWTLVMGEWMLKFYAP
WCPSCQQTDSEWEAFAKNGEILQISVGKVDVIQEPGLSGRFFVTTLPAFFHAKDGIFRRYRG
PGIFEDLQNYILEKKWQSVEPLTGWKS PASLTMMSGMAGLFSISGKIWHLHNYFTVTLGIPAW
CSYVFFVIATLVFGLFMGLVLVVI SECFYVPLPRHLSERSEQNRRSEEAHRAEQLQDAEEEK
DDSNEEENKDSLVDDEEEKEDLGDEDEAEEEEEDNLAAGVDEERSEANDQGPPGEDGVTR
EVEPEEAEEGISEQPCPADTEVVEDSLRQRKSQHADKGL
```

Important features:

Signal peptide:

amino acids 1-22

Transmembrane domain:

amino acids 191-211

N-glycosylation site.

amino acids 46-49

Thioredoxin family proteins. (homologous region to disulfide isomerase)

amino acids 56-72

Flavodoxin proteins

amino acids 173-187

FIGURE 201

ATCTGGTTGAAC TACTTAAGCTTAACTTGTAACTCCGGTAAGTACCTAGCCCACATGATT
TGACTCAGAGATTCTCTTTGTCCACAGACAGTCATCTCAGGGGCAGAAAGAAAAGAGCTCC
CAAATGCTATATCTATTCAAGGGCTCTCAAGAACAAATGGAATATCATCCTGATTAGAAAAT
TTGGATGAAGATGGATATACTCAATTACACTCGACTCTCAAAGCAATACCAGGATAGCTGT
TGTTTCAGAGAAAGGATCGTGTGCATCTCCTGGCGCTCATTGCTGTAATTGG
GAATCCTATGCTTGGTAATACTGGTGTAGCTGTGGTACCATGGGGTTCTTCC
AGCCCTTGTCTCCTAATTGGATTATATGAGAAGAGCTGTTATCTATTCAGCATGTCACT
AAATTCCCTGGGATGGAAGTAAAGACAATGCTGGCAACTGGCTCTAATCTCCTAAAGATAG
ACAGCTCAAATGAATTGGGATTATAGTAAAACAAGTGTCTCCAACCTGATAATTCAATT
TGGATAGGCCTTCTCGGCCCCAGACTGAGGTACCATGGCTCTGGGAGGATGGATCAACATT
CTCTTCTAACTTATTTCAGATCAGAACACAGCTACCCAAGAAAACCCATCTCAAATTGTG
TATGGATTCACTGTGTCAGTCATTATGACCAACTGTGTAGTGTGCCCTCATATAGTATTGT
GAGAAGAAGTTTCAATGAAGAGGAAGGGTGGAGAAGGAGAGAGAATATGTGAGGTAGTA
AGGAGGACAGAAACAGAACAGAAAAGAGTAACAGCTGAGGTCAAGATAAATGCAGAAAATG
TTTAGAGAGCTTGGCAACTGTAATCTAACCAAGAAATTGAAGGGAGAGGCTGTGATTCT
GTATTGTGACCTACAGGTAGGCTAGTATTATTTCTAGTTAGTAGATCCCTAGACATGG
AATCAGGGCAGCCAAGCTTGAGTTTATTTTATTATTATTATTATTGAGATAGGGTCT
CACTTGTGACCTGGAGTGCAGTGGCACAATCTGACTCACTGCAGCTATCTCGC
CTCAGCCCCCTCAAGTAGCTGGACTACAGGTGCATGCCACATGCCAGGCTAATTGGT
TTTTTGAGAGACTGGGTTTGCCATGTTGACCAAGCTGGCTCTAACTCCTGGCTTAAG
TGATCTGCCGCTTGGCTCCAAAGTGTGGATTACAGATGTGAGGCCACACCTGGC
CCCAAGCTGAATTTCATTCTGCCATTGACTTGGCATTACCTTGGTAAGCCATAAGCGA
ATCTTAATTCTGGCTCTACAGAGTTGTTCATGCTCAACAAATGCCATTGAAGTGCACGGT
GTGTTGCCACGATTGACCTCAACTCTAGCAGTATATCAGTTATGAACGTGAGGGTCAAAT
ATATTCTGAATAGCTAAATGAAGAAATGGAAAAATCTCACCACAGTCAGAGCAATT
ATTATTTCATCAGTATGATCATAATTATGATTATCATCTTAGTAAAAGCAGGAACCTCTA
CTTTTTCTTATCAATTAAATAGCTCAGAGAGTACATCTGCCATATCTCTAATAGAATCTT
TTTTTTTTTTTTTGAGACAGAGTTCGCTCTGTTGCCAGGCTGGAGTGCAACGG
CACGATCTGGCTCACCGCAACCTCCGCCCTGGGTTCAAGCAATTCTCCTGCCTCAGCCT
CCCAAGTAGCTGGGATTACAGTCAGGCACCACACCCGGCTAATTGTATTTTTAGT
AGAGACAGGGTTCTCCATGTCGGTAGGGTAGTCCGAACCTGACCTCAAGTGATCTGC
CTGCCTCGGCCTCCAAAGTGCTGGGATTACAGGCGTGAGCCACTGCACCCAGCCTAGAATCT
TGTATAATATGTAATTGTAGGGAAACTGCTCTCATAGGAAAGTTCTGCTTTAAATACA
AAAATACATAAAATACATAAAATCTGATGATGAATATAAAAAGTAACCAACCTCATTGGA
ACAAGTATTAACATTGGAAATATGTTTATTAGTTGTGATGTACTGTTTACAATT
ACCATTTCAGTAATTACTGTAATGGTATTATGGAATGAAACTATATTCCCTCATG
TGCTGATTGTCTTATTTCATACCTTCCACTGGTGCTATTGTTATTCCAATGGATA
TTCTGTATTACTAGGGAGGCATTACAGTCTCTAATGTTGATTAATATGTGAAAAGAAAT
TGTACCAATTACTAAATTATGCAGTTAAAATGGATGATTGTTATGTTATGTGGATTTCAT
TTCAATAAAGAAACTCTTATCAAAAAAAAAAAAAAA

FIGURE 202

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53912
<subunit 1 of 1, 201 aa, 1 stop
<MW: 22563, pI: 4.87, NX(S/T): 1
MEYHPDLENLDEDGYTQLHFDSQSNTRIAVVSEKGSCAASPPWRLIAVILGILCLVILVIAV
VLGTMGVLSLSSPCPPNIIYEKSCYLFMSMSLNSWDGSKRQCWQLGSNLLKIDSSNELGFIVKQ
VSSQPDNSFWIGLSPQTEVPWLWEDGSTFSSNLFQIRTTATQENPSPNCVWIHVSIVYDQL
CSVPSYSICEKKFSM
```

Important features:

Type II transmembrane domain:

amino acids 45-65

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 197-200

N-myristoylation sites.

amino acids 35-40 and 151-156

Homologous region to LDL receptor

amino acids 34-67 and 70-200.

FIGURE 203

FIGURE 204

```
</usr/seqdb2/sst/DNA/Dnaseqs,min/ss.DNA50921
<subunit 1 of 1, 693 aa, 1 stop
<MW: 77738, pI: 8.87, NX(S/T): 7
MTPQSLLQTTLFLLSLLFLVQGAHGRGHREDFRFCQRNQTHRSSLHYKPTPDLRISIENSE
EALTVHAPFPAAHPASRSFPDPRLYHFCLYWNRHAGRLHLLYGRDFLLSDKASSLLCFQH
QEESLAQGPPLLATSVTSWWSPQNISLPSAASFTFSFHSPPHTAAHNASVDMCELKRDQLL
SQFLKHPQKASRRPSAAPASQQLQSLSKLTSVRFMGDMVSFEEDRINATVWKLQPTAGLQD
LHIHSRQEEEQSEIMEYSVLLPRTLFQRTKGRSGAEKRLLVDFSSQALFQDKNSSQVLGE
KVLGIVVQNTKVANLTEPVVLTFQHQLQPKNVTLQCVFWVEDPTLSSPGHSSAGCETVRRE
TQTSCFCNHLYTFAVLMVSSVEVDAVHKHYLSLLSYVGCVVSALA CLVTIAAYLCSRVPPLC
RRKPRDYTIKVHMNLLAVFLLDTSFLLSEPVALTGSEAGCRASAIFLHFSSLTCLSWMGLE
GYNLYRLVVEVFGTYVPGYLLKLSAMGWGFPIFLVTVALVDVDNYGPIILAVHRTPEGVIY
PSMCWIRDSLVSYITNLGLFSLVFLFNAMMLATMVVQILRLRPHTQKWSHVLTLLGLSLVLG
LPWALIFFSFASGTFQLVVLYLFSIITSFQGFLIFIWYWSMRLQARGGPSPLKNSDSARLP
ISSGSTSSSRI
```

Important features:

Signal peptide:

amino acids 1-25

Putative transmembrane domains:

amino acids 382-398, 402-420, 445-468, 473-491, 519-537, 568-590
and 634-657

Microbodies C-terminal targeting signal.

amino acids 691-693

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 198-201 and 370-373

N-glycosylation sites.

amino acids 39-42, 148-151, 171-174, 234-237, 303-306, 324-327
and 341-344

G-protein coupled receptors family 2 proteins

amino acids 475-504

FIGURE 205

TGCCTGGCCTGCCTGTCAACAATGCCGCTACTCTGCTTCCAGGTTGCCCTGCCTGCAGA
GGAAANCNTCGGGACTACACCNNTCAAGTGCACATGAACCTGCTGCTGGCCGTCTCCTGCTG
GACACGAGCTTCCTGCTCAGCGNAGCCGGTGGCCCTGACAGGCTCTGAAGGCTGGCTGCCGA
GCCAGTGCCATCTTCCTGCACTTCTCCTGCTCACCTGCCTTCCTGGATGGCCTCGAGGGGG
TACAACCTCTACCGACTCGTGGTGGAGGTCTTGGCACCTATGTCCCTGGCTACCTACTCAA
GCTGAGGCCATGGCTGGGCTTCCCCATCTTCTGGTGACGCTGGTGGCCCTGGTGGATG
TGGACAACATGGCCCCATCATCTTGGCTGTGCATAGGACTCCAGAGGGCGTCATCTACCCT
TCCATGTGCTGGATCCGGACTCCCTGGTCAGCTACATCACCAACCTGGCCTTCAACGCT
GGTGTTCCTGTTAACATGG

FIGURE 206

CGGACGCGTGGCGGACCGTGGCGGACCGTGGCGGACCGTGGCGTGGCTGGTCAGGTCCAGGTTGCTTGA
TCCTTTCAAAAAGTGGAGACACAGAAGAGGGCTCTAGGAAAAGTTGGATGGATTATGTGAAACTACCC
GCGATTCTCGCTGCCAGAGCAGGCTCGGCCTCCACCCAGTCAGCAGCCATTCCCGTGGTGAAGAGAC
TCGGGAGTCGCTGCTCCAAAGTGGCCCGTGAAGCTCTCACCCAGTCAGCCAA**ATG**AGCCTCTCGGC
TTCTCCTGCTGACATCTGCCCTGGCGGAGAGACAGGGACTCAGGCGGAATCCAACCTGAGTAGTAAATTCC
AGTTTCCAGCAACAAGGAACAGAACGGAGTACAAGATCCTCAGCATGAGAGAATTATTACTGTGTCTACTAATG
GAAGTATTACAGCCCAGGTTCTCTCATACTTATCAAGAAATACGGTCTTGGTATGGAGATTAGTAGCAGTAG
AGGAAAATGTATGGATACAACCTACGTTGATGAAAGATTTGGCTTGAAGACCCAGAAGAGATGACATATGCAAGT
ATGATTTGAGAGTGGAGAACCCAGTGATGGAACATATATTAGGCGCTGGTGTGGTACTGTACCG
GAAAACAGATTCTAAAGGAATCAAATTAGGATAAGATTGTATCTGATGAATATTTCCTCTGAACCAAGGGT
TCTGCATCCACTACAACATTGTCATGCCACAATTACAGAAGCTGTGAGTCCTCAGTGTACCCCTCAGCTT
TGCCACTGGACCTGCTTAATAATGCTATAACTGCCTTAGTACCTTGAAGACCTTATTGATATCTGAACAG
AGAGATGGCAGTTGGACTTAGAAGATCTATATAGGCCAACTTGGCAACTTCTGGCAAGGCTTTGGAA
GAAAATCCAGAGTGGTGGATCTGAACCTCTAACAGAGGGAGTAAGATTATACAGCTGCACACCTCGTAACCTCT
CAGTGTCCATAAGGGAAACTAAAGAGAACGATACCATTCTGGCCAGGTTGTCCTGGTAAACGCTGTG
GTGGGAACTGTGCCTGTTGCTCCACAATTGCAATGAATGTCAATGTGTCCAAGCAGAAAGTTACTAAAAAATACC
ACGAGGTCTTCAGTTGAGACCAAAGACGGTGTCAAGGGATTGACAAATCACTCACCAGCAGTGGCCCTGGAGC
ACCATGAGGAGTGTGACTGTGTGCAAGAGGGAGCACAGGAGG**ATG**CCGATCACCACCAGCAGCTTGGCCA
GAGCTGTGCAAGTGCAGTGGCTGATTCTATTAGAGAACGTATGCGTTATCTCCATCCTTAATCTCAGTTGTTGCT
TCAAGGACCTTCATCTCAGGATTACAGTCATTGAAAGAGGAGACATCAAACAGAATTAGGAGTTGTGCA
ACAGCTCTTGAGAGGAGGCCTAAAGGACAGGAGAAAAGGTCTCAATCGTGGAAAGAAAATTAAATGTGTAT
TAAATAGATCACCAGCTAGTTGAGTTACCATGTCAGTATTCCACTAGCTGGTTCTGTATTTCAGTTCTTC
GATACGGCTTAGGTAATGTCAGTACAGGAAAAAAACTGTGCAAGTGAGCACCTGATTCCGTTGCTTGAAC
TCTAAAGCTCATGTCCTGGCTAAATCGTATAAAATCTGGATTTTTTTTGCTCATATTACACAT
ATGTAACACCAGAACATTCTATGACTACAAACCTGGTTTAAAGGAACATGTTGCTATGAATTAAACTTGT
GTCATGCTGATAGGACAGACTGGATTTCATATTCTTATTAAATTCGCCATTAGAAGAAGAGAACATACA
TTCATGGTTGAGAGATAAACCTGAAAAGAAGACTGGCCTTACCTTATCGATAAGTCAGTTATTG
TTCATGGTGTACATTATATTCTCCTTGTACATTATAACTGTTGGCTTTCTAATCTGTTAAATATATCT
ATTTCATGAAAGGTATTAAATTCTTTATGACAACCTAGTCAACTATTTCAGCTTGTAAATTTC
AAACACAATTGTTAGCCAGAGGAACAAAGATGATAAAATATTGTTGCTCTGACAAAATACATGTATTCA
TTCTCGTATGGTGTAGAGTTAGATTACCTGCAATTAAACTGAAATTGGAATAGAATTGTAAGTTGCAAA
GACTTTTGGAAAATAATTAAATTATCATATCTCCATTCTGTATTGGAGATGAAAATAAAAGCAACTTATGA
AAGTAGACATTAGATCCAGGCCATTACTAACCTATTCTTTGGGAAACTGAGCCTAGCTCAGAAAACAT
AAAGCACCTGAAAAGACTTGGCAGCTCCTGATAAAAGCGTGTGCTGTGAGTAGGAACACATCCTATT
TTGTGATGTTGTGGTTTATTATCTTAAACTCTGTTCCATACACTGTATAAAATACATGGATATTGTTATGTACA
GAAGTATGTCCTTAACCAGTTACTTATTGACTCTGGCAATTAAAAGAAAATCAGTAAAATATTGCTTGT
AAAATGCTTAATATNGTGCCTAGGTTATGTGGTACTATTGAATCAAAATGTATTGAATCATCAAATAAAGA
ATGTGGCTATTGGGAGAAAATTAAAAAAAGGTTAGGGATAACAGGGTAATGCGGCC

FIGURE 207

MSLFGLLLTSLAGQRQGTQAESNLSSKFQFSSNKEQNGVQDPQHERIITVSTNGSIHS
FPHTYPRNTVLVWRLVAVEENVWIQLTFDERFGLEDPEDDICKYDFVEVEEPSDGTILGRWC
GSGTVPGKQISKGNQIRIRFVSDYEFPSEPGFCIHYNIVMPQFTEAVSPSVLPPSALPLDLL
NNAITAFSTLEDLIRYLEPERWQLDLEDLYRPTWQLLGKAFVFGRKSRVVDLNLLTEEVRLY
SCTPRNFSVSIREELKRTDTIFWPGCLLVKRCGGNCACCLHNCNECQCVP SKVTKYHEVLQ
LRPKTGVRGLHKSLTDVALEHHEECDCVCRGSTGG

Signal sequence:

amino acids 1-14

FIGURE 208

CCCATCTCAAGCTGATCTGGCACCTCTCATGCTCTGCTCTCAACCAGACCTTACATTCCATTGGAAAGA
AGACTAAAAATGGTGTTCATGGACACTGAAGAGACAAATTCTTACACATAATCCTAATTG
AAACTCCTGGGGCTAGATGGTTCCCTAAACTCTGCCCTGTGATGTCACTCTGGATGTTCAAAGAACCATG
ATCGTGGACTGCACAGACAAGCATTTGACAGAAATTCTGGAGGTATTCCCACGAACACCACGAACCTCACCC
ACCATTAACCACATACCCAGACATCTCCCAGCCTTACAGACTGGGACATCTGGTAGAGATGATTTCAGA
TGCAACTGTGTACCTATTCCACTGGGGTCAAAAAACAAACATGTGATCAAGAGGCTGAGATTAAACCCAGAAGC
TTAGTGGACTCACTTAAACCTTACCTGGGAAACCGACTAGAGATAACCGCAGGGCTCCCG
CCTAGCTTACAGCTCTCAGCCTGAGGCCAACACATCTTCCATCAGAAAAGAGAATCTAACAGAACTGGCC
AACATAGAAATACTCTACCTGGGCAAAACTGTTATTATCGAAATCCTGTTATGTTCATATTCAATAGAGAAA
GATGCCCTCTAAACTGACAAAGTTAAAGTGTCTCCCTGAAAGATAACAAATGTACAGCCGCTCCACTGTT
TTGCCATCTACTTAAACAGAACTATATCTACAAACAACATGATTGCAAAATCCAAGAAGATGATTAAATAAC
CTCAACCAATTACAAATTCTGACCTAAGTGGAAATTGCCCTGTTATAATGCCCAATTCTGTGCGCCG
TGTAAAATAATTCTCCCTACAGATCCCTGTAATGCTTTGATGCGCTGACAGAAATTAAAGTTTACGTCTA
CACAGTAACTCTCTTCAGCATGTGCCCCAAGATGGTTAAGAACATCAACAAACTCCAGGAACGGATCTGTCC
CAAAACTCTGGCCAAAGAAATTGGGATGCTAAATTCTGCATTTCTCCCTGAGCCTCATCCAATTGGATCTG
TCTTCAATTGAACTTCAGGCTATCGGCATCTGAATCTATCACAAGCATTCTTCACTGAAAAGCCTG
AAAATTCTGCGGATCAGAGGATATGTTAAAGAGTTAAAGCTTAACTCTGCCATTACATAATCTCAA
AATCTTGAAGTTCTGATCTGGCACTAACCTTATAAAATTGCTAACCTCAGCATGTTAAACAATTAAAGA
CTGAAAAGTCATAGATCTTCAGTGAATAAAATCACCTCAGGAGATTCAAGTGAAGTTGGCTCTGCTCAAAT
GCCAGAACTCTGAGAAAGTTATGAAACCCAGGTCTGGAACAAATTACATTATTCAGATATGATAAGTATGCA
AGGAGTTGCGAGATTCAAAAACAAGAGGCTTCTTCATGTCGTTAATGAAAGCTGCTACAAGTATGGCAGACC
TTGGATCTAAGTAAAATAGTATTTTGTCAAGTCCCTGATTTCAGCATTCTCTTCCCAAATGCC
AATCTGTCAGGAAATCTCATTAGCCAACCTTAAATGGCAGTGAATTCCAACCTTACGAGAGCTGAGATATTG
GACTCTCCAACAAACCCGGCTGATTACTCCATTCAACAGCATTGAAGAGCTTCACAAACTGGAAGTTCTGGAT
ATAAGCAGTAATAGCATTATTCAATCAGAGGAATTACTCATATGCTAAACTTACCAAGAACCTAAAGGTT
CTGAGAAACTGATGATGAACGACAATGACATCTTCCCTCCACAGCAGGACATGGAGAGTGTGAGTCTTCTAGA
ACTCTGGAAATTAGAGGAAATCACTTAGATGTTTATGGAGAGAAGGTGATAACAGATACTTACAATTATTCAAG
AATCTGCTAAAATTAGAGGAAATTAGACATCTCTAAATTCCCTAAGTTCTGCCTCTGGAGTTTGATGGT
ATGCCTCCAAATCTAAAGAATCTCTTGGCAAAATGGGCTCAAATCTTCAGTGGAAAGAAACTCCAGTGT
CTAAAGAACCTGGAAACTTGGACCTCAGGCCACAACCAACTGACCAACTGTCCTGAGAGATTATCCAACCTGTTCC
AGAACGCTCAAGAATCTGATTCTTAAGAATAATCAAATCAGGAGTCTGACGAAGTATTTCTACAAGATGCC
CAGTTGCGATATCTGGATCTCAGCTCAAATAAAATCCAGATGATCCAAAAGACCAGCTCCAGAAAATGCC
AACAACTCTGAAGATGTTGCTTTGCATCATAATGGTTCTGTCACCTGTGATGCTGTTGTTGCTGGTGG
GTTAACCATACGGAGGTGACTATTCTTACCTGGCCACAGATGTGACTTGTGTTGGGGCAGGAGCACACAAGGGC
CAAAGTGTGATCTCCCTGGATCTGTACACCTGTGAGTTAGATCTGACTAACCTGATTCTGTTCTCACTTCCATA
TCTGTATCTCTCTTCTCATGGTGTGATGACAGCAAGTCACCTCTATTCTGGGATGTTGTTGTTGATATTACCAT
TTCTGTAAGCCAAGATAAAGGGTATCAGCGTCTAATATCACCAGACTGTTGCTATGATGCTTTATTGTT
GACACTAAAGACCCAGCTGTGACCGAGTGGGTTGGCTGAGCTGGTGGCCTAACTGGAAGACCCAAAGAGAGAAA
CATTTAAATTATGTCGAGGAAAGGGACTGGTACCCAGGGCAGCCAGTCTGGAAAACCTTCCCAGAGCATA
CAGCTTAGCAAAAAGACAGTGTGTTGTGATGACAGACAAGTATGCAAAGACTGAAAATTAAAGATAGCATTTCAC
TTGTCCTCCAGGGCTCATGGATGAAAAAGTGTGATTATCTGATATTCTGAGAAGGCCCTTCAGAAG
TCCAAGTTCTCCAGCTCCGGAAAAGGCTCTGTGGGAGTTCTGTCCTGAGTGGCCAACAAACCCGCAAGCTCAC
CCATACTCTGGCAGTGTCTAAAGAACGCCCTGGCCACAGACAATCATGTGGCCTATAGTCAGGTGTTCAAGGAA
ACGGTCTAGCCCTTGTCAAAACACAACCTGCCAGTTACCAAGGAGAGGCC
GGC

FIGURE 209

MVFPMWTLKRQILILFNIILISKLLGARWFPKTLPCDVTLDPKHNHIVDCTDKHLTEIPGG
IPTNTTNLTLTINHIPDISPASFHRLDHLEIDFRNCVPILGSKNNMCIKRLQIKPRSFS
GLTYLKSLYLDGNQLEIPQGLPPSLQLLSLEANNIFSIRKENTELANIEILYLGQNCYYR
NPCYVSYSIEKDAFLNLTKLVLSLKDNNVTAVPTVLPSTLTELVLYNNMIAKIQEDDFNNL
NQLQILDLSGNCPRCYNAPFPCAPCKNNSPLOIPVNAFDALTELKVLRLHSNSLQHVPPRWF
KNINKLQELDLSQNFLAKEIGDAKFLHFLPSLIQLDLSFNFELQVYRASMNLSQAFSSLKSL
KILRIRGYVFKELKSFNLSPFHNLQNLEVLDLGTNFKIANLSMFQFKRLKVIDLSVNKIS
PSGDSSEVGFCNSARTSVESYEPQVLEQLHYFRYDKYARSCRFKNKEASFMSVNESCYKGQ
TLDLSKNSIFFVKSSDFQHLSFLKCLNLSGNLISQTLNGSEFQPLAELRYLDFSNRNL
STA FEELHKLEVLDI SSNSHYFQSEGITHMLNFTKNLKVLQKLMNDNDI SSSTSRTMESES
LRTLEFRGNHLDVLWREGDNRYLQLFKNLLKLEELDISKNSLSFLPSGVFDGMPPNLKNL
AKNGLKSFWSKKLQCLKNLETLDLSHNQLTTVPERLSNCSRSLKNLILKNNQIRSLTKYFLQ
DAFQLRYLDLSSNKIQMIQKTSFPENVLNNLKMLLHHNRFLCTCDAVWFVWWVNHTEV
YIYHFCKAKIKGYQRLISPDCCYDAFIVYDTKDPAVTEWVLAELVAKLEDPREKHFNL
RDWLPQPVLENLSQSIQLSKKTVFVMTDKYAKTENFKIAFYLSHQRLMDEKV
KPFQKSFKFLQLRKRLCGSSVLEWPTNPQAHPYFWQCLKNALATDNHV
AYSQVFKETV

Signal sequence:

amino acids 1-26

Transmembrane domain:

amino acids 840-860

FIGURE 210

FIGURE 211

MENMFLQSSMLTCIFLLISGSCELCAEENFSRSYPDEKKQNDSVIAECNRRLQEVPQTGV
KYVTELDLSDNFITHITNESFQGLQNLTKINLNHNPNVQHQNQNPQIQSNGLNITDGAFLN
KNLRELLLEDNQLPQIPSGLPESLTELSSIQNNIYNITKEGISRLINLKNLYLAWNCFNK
CEKTNIEDGVFETLTNLELLSLSFNLSHVPPKLPSSLRKLFLSNTQIKYISEEDFKGLIN
TLDDLSGNCPRCFNAPFPCVPCDGGASINIDRFAFQNLTQLRYLNLSSTSLRKINAAWFKNM
PHLKVLDFNYLVGEIVSGAFLTMLPRLEILDLSFNYIKGSYPQHINISRNFSKLLSLRAL
HLRGYVFQELREDDFQPLMQLPNLSTINLGINFQIDFKLFQNFNSNLEIIYLSERISPLV
KDTRQSYANSSSFQRHIRKRRSTDFFDPHSNFYHFTRPLIKPQCAAYGKALDLSLNSIFFI
GPNQFENLPDIACLNLSANSNAQVLSGTEFSAI PHVKYLDLTNNRLDFDNASALTELDLEV
LDLSYNSHYFRIAGVTHHLEFIQNFTNLKVLNLSHNNIYTLDKYNLESKSLVELVFSGNRL
DILWNDDDNRYISIFKGLKNLTRLDSLNRKHIPNEAFLNLPASLTELHINDNMLKFFNWT
LLQQFPRLELLDLRGNKLLFLTDSSLSDFTSSLRTLLSHNRISHLPSGFLSEVSSLKHLDS
SNLLKTINKSALETKTTKLSMLELHGNPFECTCDIGDFRRWMDEHLNVKIPRLVDVICASP
GDQRGKSIVSLELTTCVSDVTAVILFFFITTMVMLAALAHHLFYWDVWFIVNVCLAKVK
GYRSLSTSQTFYDAYISYDTKDASVTDWVINELRYHLEESRDKNVLLCLEERDWDPGLAIID
NLMQSINQSKKTVFVLTKKYAKSWNFKTAFYLALQRLMDENMDVIIFILEPVLQHSQYLR
RQRICKSSILQWPDPNPKAEGLFWQTLRNVVLTENDSRYNNMYVDSIKQY

Signal sequence:

amino acids 1-26

Transmembrane domain:

amino acids 826-848

FIGURE 212

CCAGGTCCAAC TGCACCTCGGTTCTATCGATTGAATTCCCCGGGGATCCTCTAGAGATCCCTCGACCTCGACCCACGCGTCCGCCAAGCTGGCCCTGCACGGCTGCAAGGGAGGCTCCTGTGGA CAGGCCAGGCAGGTGGCCTCAGGAGGTGCCTCAGGCCAGTGGCCTGAGGCCAGC AAGGGCTAGGGTCCATCTCCAGTCCCAGGACACAGCAGCGGCCACCATGCCACGCCCTGGCTCCAGCAGCATCAGCAGCCCCAGGACCGGGAGGCACAGGTGGCCCCACCACCCGGAGGA GCAGCTCCTGCCCTGTCCGGGGATGACTGATTCTCCTCCGCCAGGCCACCCAGAGGAGAA GGCCACCCCGCCTGGAGGCACAGGCCATGGGGCTCTCAGGAGGTGCTGCTGATGTGGCTT CTGGTGTGGCAGTGGCGGCACAGAGCACGCCTACCGGCCGGCGTAGGGTGTGTGCTGT CGGGGCTCACGGGACCCCTGTCTCCGAGTCGTTCGCAGCGTGTGTACCAGCCCTCCTCA CCACCTGCGACGGCACCGGCCCTGCAGCACCTACCGAACCATCTATAGGACCGCCTACCGC CGCAGCCCTGGCTGGCCCTGCCAGGCCTCGCTACCGTGTGTGCCCGGCTGGAAGAGGAC CAGCAGGGCTTCCGGGCCTGTGGAGCAGCAATATGCCAGCCCATGCCGGAACCGGAGGGAG GCTGTGTCCAGCCTGGCGCTGCCGCTGCCCTGCAGGATGGGGGTGACACTGCCAGTCA GATGTGGATGAATGCAGTGCTAGGAGGGCGGCTGTCCCCAGCGCTGCATCACACCGCCGG CAGTTACTGGTGCAGTGTGGAGGGCACAGCCTGTCTGCAGACGGTACACTCTGTGTGC CCAAGGGAGGGCCCCCAGGGTGGCCCCAACCGACAGGAGTGGACAGTGAATGAAGGAA GAAGTGCAGAGGCTGCAGTCCAGGGTGGACCTGCTGGAGGAGAAGCTGCAGCTGGTGTGGC CCCACTGCACAGCCTGCCCTCGCAGGCAGTGGAGCATGGCTCCGGACCCGGCAGCCTCC TGGTGCACTCCTTCCAGCAGCTGGCCGCATCGACTCCCTGAGCGAGCAGATTCCCTCCTG GAGGAGCAGCTGGGTCTGCTCCTGCAAGAAAGACTCGTGACTGCCAGCGCCCAGGCTG GACTGAGCCCTCACGCCGCCCTGCAGCCCCATGCCCTGCCAACATGCTGGGGTCCAG AAGCCACCTGGGGTGAUTGAGCGGAAGGCCAGGCAGGGCCTCCTCCTCTCCCTCCCC TTCCTCGGGAGGCTCCCCAGACCCCTGGCATGGATGGCTGGATCTTCTGTGAATCCAC CCCTGGCTACCCCCACCCCTGGCTACCCCAACGGCATCCAAAGGCCAGGTGGCCCTCAGCTG AGGGAAAGGTACGAGCTCCCTGCTGGAGCCTGGACCCATGGCACAGGCCAGGCAGCCGGAG GCTGGGTGGGGCCTCAGTGGGGCTGCTGCCCTGACCCCCAGCACAATAAAATGAAACGTGA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGGCCGCGACTCTAGAGT CGACCTGCAGAAGCTTGGCCGCATGGCCAACTTGTTATTGCAGCTATAATGGTTACAAAT

FIGURE 213

MRGSQEVL LMWLLV LAVGGTEHAYRPGRRVCAVRAHGD PVSE SFVQRVYQPFLTTCDGHRAC
STYRTIYRTAYRRSPGLAPARPRYACCPGWKRTSGLPGACGAAICQPPCRNGGSCVQPGRCR
CPAGWRGDTCQSDVDECSARRGGCPQRCINTAGSYWCQCWE GHSL SADGTLCVPKGGPPRVA
PNPTGVDSAMKEEVQRLQSRVDLLEEKLQLVLAPLHSLASQALEHGLPDPGSLLVHSFQQLG
RIDSLSEQISFLEEQLGSCSCKKDS

Signal sequence:

1-19

FIGURE 214

GCCAGGCAGGTGGCCTCAGGAGGTGCCTCAGGCAGGCCAGTGGCCTGAGGCCAGCAAG
GGCTAGGGTCCATCTCCAGTCCCAGGACACAGCAGCGGCCACCATGCCACGCCCTGGCTCC
AGCAGCATCAGAGCAGCCCTGTGGTGGCAGCAAAGTCAGCTTGGCTGGCCCGCTGTGA
GGGGCTTCGCGCTACGCCCTGCGGTGTCCCAGGGCTGAGGTCTCCTCATCTCTCCCTAGC
AGTGGATGAGCAACCCAACGGGGCCGGGGAGGGAACTGGCCCCGAGGGAGAGGAACCCC
AAAGCCACATCTGTAGCCAGGATGAGCAGTGTGAATCCAGGCAGCCCCCAGGACCAGGGAGG
CACAGGTGGCCCCCACCACCCGGAGGAGCAGCTCCTGCCCTGTCCGGGGATGACTGATT
TCCTCCGCCAGGCCACCCAGAGGAGAAGGCCACCCGCCCTGGAGGCACAGGCCATGAGGGC
TCTCAGGAGGTGCTGCTGATGTGGCTTCTGGTGTGGCAGTGGGGGGCACAGAGCACGCC
CCGGCCCGGCCGTAGGGTGTGTGCTGCCGGCTCACGGGACCCCTGTCTCCGAGTCGTT
TGCAGCGTGTGTACCAAGCCCTCCTCACCACTGCAGGGCACGGGCCTGCAGCACCTAC
CGAACCATCTATAAGGACCGCCTACCGCCGAGCCCTGGCTGGCCCTGCCAGGCCTCGCTA
CGCGTGCTGCCCGGCTGGAAGAGGACCAGCGGGCTTCCTGGGCTGTGGAGCAGCAATAT
GCCAGCCGCCATGCCGGAACGGAGGGAGCTGTGTCCAGCCTGGCCGTGCCGTGCC
GGATGGCGGGGTGACACTTGCAGTCAGATGTGGATGAATGCAGTGCTAGGAGGGCGGCTG
TCCCCAGCGCTGCATCAACACCGCCGGCAGTTACTGGTGCAGTGTGGAGGGCACAGCC
TGTCTGCAGACGGTACACTCTGTGTGCCAAGGGAGGGCCCCCAGGGTGGCCCCAACCCG
ACAGGAGTGGACAGTGCAATGAAGGAAGAAGTGCAGAGGCTGCAGTCAGGGTGGACCTG
GGAGGAGAAGCTGCAGCTGGTGTGGCCCCACTGCACAGCCTGGCCTGCAGGCAGTGGAGC
ATGGGCTCCGGACCCCGGCAGCCTGGTGCACTCCTCCAGCAGCTGGCCGCATCGAC
TCCCTGAGCGAGCAGATTCCTCCTGGAGGGAGCAGCTGGGCTCTGCTCTGCAAGAAAGA
CTCGTGACTGCCAGCGCTCCAGGCTGGACTGAGCCCTCACGCCGCCCTGCAGCCCCATG
CCCCTGCCAACATGCTGGGGTCCAGAACGCCACCTGGGACTGAGCGGAAGGCCAGGC
AGGGCCTTCCTCCTCTCCCTCCCTGGGAGGCTCCCCAGACCCTGGCATGGGAT
GGGCTGGGATCTCTGTGAATCCACCCCTGGCTACCCCAACCCCTGGCTACCCCAACGG
TCCCAAGGCCAGGTGGACCCCTCAGCTGAGGGAAAGGTACGAGCTCCCTGCTGGAGCCTGG
CCATGGCACAGGCCAGGCAGCCGGAGGCTGGTGGGCCTCAGTGGGGCTGCTGCCTGAC
CCCCAGCACAATAAAATGAAACGTG

FIGURE 215

MRGSQEVL LMWLLV LAVGGTEHAYRPGR RVCAVRAHGD PVSE SFVQRV YQPFLTTCDGHRAC
STYRTIYRTAYRRSPGLAPARPRYACCPG WKRTSGLPGACGAAICQPPCRNGGSCVQPGRCR
CPAGWRG DTCQSDVDECSARRGGCPQRCINTAGSYWCQCWE GHSL SADGTL CVPKG GPPRVA
PNPTGVDSAMKEEVQRLQSRVDLLEEKLQLVLAPLHSLASQALEHGLPD PGSLLVHSFQQLG
RIDSLSEQISFLEEQLGSCSCKKDS

Signal sequence:

1-19

69
68
67
66
65
64
63
62
61
60
59
58
57
56
55
54
53
52
51
50
49
48
47
46
45
44
43
42
41
40
39
38
37
36
35
34
33
32
31
30
29
28
27
26
25
24
23
22
21
20
19
18
17
16
15
14
13
12
11
10
9
8
7
6
5
4
3
2
1

FIGURE 216

CCCACGCGTCCGAAGCTGGCCCTGCACGGCTGCAAGGGAGGCTCCTGTGGACAGGCCAGGCA
GGTGGGCCTCAGGAGGTGCCTCCAGGCAGGCCAGTGGCCTGAGGCCAGCAAGGGCTAGGG
TCCATCTCCAGTCCCAGGACACAGCAGCGGCCACCATGCCACGCCCTGGCCTCAGCAGCAT
CAGCAGCCCCAGGACCGGGGAGGCACAGGTGGCCCCACCACCCGGAGGAGCAGCTCCTGC
CCCTGTCCGGGGATGACTGATTCTCCTCCGCCAGGCCACCCAGAGGAGAAGGCCACCCGC
CTGGAGGCACAGGCCATGAGGGCTCTCAGGAGGTGCTGCTGATGTGGCTTCTGGTGTGGC
AGTGGGCGGCACAGAGCACGCCACCGGCCGGCGTAGGGTGTGCTGTCCGGCTCACG
GGGACCCCTGTCTCGAGTCGTTCGCAGCGTGTGTACCGCCCTTCCTCACCACTGCGAC
GGGCACCGGGCCTGCAGCACCTACCGAACCATCTATAGGACCGCTACCGCCGCAGCCCTGG
GCTGGCCCTGCCAGGCCTCGCTACCGTGTGCCCGGCTGGAAGAGGACCAGCGGGCTTC
CTGGGGCCTGTGGAGCAGCAATATGCCAGCCCATGCCGGAACGGAGGGAGCTGTGTCCAG
CCTGGCCGCTGCCGCTGCCCTGCAGGATGGCGGGGTGACACTTGCCAGTCAGATGTGGATGA
ATGCAGTGCTAGGAGGGCGGCTGTCCCCAGCGCTGCGTCAACACCGCCGGCAGTTACTGGT
GCCAGTGTTGGAGGGCACAGCCTGTCTGCAGACGGTACACTCTGTGTGCCAACGGGAGGG
CCCCCCAGGGTGGCCCCCAACCGACAGGAGTGGACAGTGCAATGAAGGAAGAAGTGCAGAG
GCTGCAGTCCAGGGTGGACCTGCTGGAGGAGAAGCTGCAGCTGGTGTGCCCAAGGGAGGG
GCCTGGCCTCGCAGGCACTGGAGCATGGCTCCGGACCCGGCAGCCTCTGGTGCACTCC
TTCCAGCAGCTCGGCCGCATCGACTCCCTGAGCGAGCAGATTCTTCTGGAGGAGCAGCT
GGGTCCTGCTCTGCAAGAAAGACTCGTGACTGCCAGGCCAGGCTGGACTGAGCCCC
TCACGCCGCCTGCAGCCCCATGCCCTGCCAACATGCTGGGGTCCAGAACGCCACCTCG
GGGTGACTGAGCGGAAGGCCAGGCAGGGCTTCCCTCTCTCCCTCCCTCGGGAG
GCTCCCCAGACCTGGCATGGATGGCTGGATCTTCTCTGTGAATCCACCCCTGGCTACC
CCCACCCCTGGCTACCCCAACGGCATCCAAAGGCCAGGTGGCCCTCAGCTGAGGGAAAGGTAC
GAGCTCCCTGCTGGAGCCTGGACCCATGGCACAGGCCAGGCAGCCGGAGGCTGGTGGGG
CCTCAGTGGGGCTGCTGCCTGACCCCCAGCACAATAAAATGAAACGTG

FIGURE 217

MRGSQEVL LMWLLV LAVGGTEHAYRPGRRVCAVRAHGD PVSE SFVQRVYQPFLTTCDGHRAC
STYRTIYRTAYRRSPGLAPARPRYACCPGWKRTSGLPGACGAAICQPPCRNGGSCVQPGRCR
CPAGWRGDTCQSDVDECSARRGGCPQRCVNTAGSYWCQCWE GHSL SADGTLCVPKGGPPRVA
PNPTGVDSAMKEEVQRLQSRVDLLEEKLQLVLAPLHSLASQALEHGLPD PGSLLVHSFQQLG
RIDSLSEQISFLEEQLGSCSCKKDS

Signal sequence:

1-19

0964333333 0964333333 0964333333

FIGURE 218

GGTTGCCACAGCTGGTTAGGGCCCCGACCACGGGGCCCTTGTCAAGGAGGAGACAGCCTCCGGCCGGGGAG
GACAAGTCGCTGCCACCTTGGCTGCCGACGTGATTCCCTGGACGGTCCGTTCTGCCGTCAAGCTGCCGGCG
AGTTGGGCTCCCGTGTTCAGGCCGCTCCCCCTTCTGGTCTCCCTCTCCCGCTGGGGGTTATCGGGAGG
AGATTGTCTCCAGGGCTAGCAATTGGACTTTGATGATGTTGACCCAGCAGGAAATAGCAGGCAACGTGAT
TTCAAAGCTGGGCTCAGCCTCTGTTCTCTCGTCAATCGAAAACCCATTGGAGCAGGAATTCCAATCA
TGTCTGTGATGGTGGTGAGAAAGAAGGTGACACGGAAATGGGAGAAACTCCCAGGAGGAACACCTTTGCTGTG
ATGGCCGCGTCATGAGGCCGAAAAGGGCATTTCTACCTGACCCCTTCTCATCCTGGGGACATGTACAC
TCTTCTCGCCCTTGAGTGGCGTACCTGGCTTCACTGTCCTGCCATCCCTGATTGCTGCCATGCTCT
TCCTTCTCCATGGCTACACTGGAGGACAGCTTCACTGACCCCTGGAGTGAATTCTCGGGCGTACCAAGATG
AAGCAGCTTCTCATAGAAATGGAGATAGAAGCTACCAATGGTGGGTGCCCCAGGGCAGCAGCACCACGCC
TCAAGAAATTCCAGATAAACAAACAGATTGTGAAACTGAAATACTGTTACACATGCAAGATCTTCCGGCTCCCC
GGGCCTCCCATTGCAAGCATCTGTGACAACACTGTGAGCGCTCGACCATCACTGCCCTGGGGGGAAATTGTG
TTGGAAAGAGGAACCTACCGCTACTTCTACCTCTCATCCTTCTCTCCCTCCTCACAAATCTATGTTCTCGCT
TCAACATGTCATGGCCCTCAAATCTTGGAAAATTGGCTTCTGGAGACATTGAAAGAAACTCCTGGAAC
TTCTAGAAGTCCCTCATTGCTTCTTACACTCTGGTCCGTCGTGGACTGACTGGATTCTACACTTCTCGTGG
CTCTCAACCAGACAACCAATGAAAGACATCAAAGGATCATGGACAGGGAGAATCGCTCCAGAACCTACAGCC
ATGGCAATATTGTGAAGAAACTGCTGTGAAGTGCTGTGGCCCTTGCCCCCAGTGTGCTGGATGAAAGGGTA
TTTGCCACTGGAGGAAAGTGGAGTCGACCTCCCAGTACTCAAGAGACCAGTAGCAGCCTCTGCCACAGAGCC
CAGCCCCACAGAACACCTGAACCAAATGAGATGCCGGAGGACAGCAGCACTCCGAAGAGATGCCACCTCCAG
AGCCCCCAGAGCCACACAGGAGGAGCTGAAGTAGCCTATCTATGGAAGAGACTTTGTTGTGTT
TAATTAGGGCTATGAGAGATTCAAGGTGAGAAGTTAAACCTGAGACAGAGAGCAAGTAAGCTGTCCTTTAACT
GTTTTCTTGGCTTTAGTCACCCAGTTGCACACTGGCATTTCTGTCAGCTTTAAATTCTGA
CAAGGCAGTGGCAGAAAGATGTCAGTCACCTCTGATAACTGGAAAATGGGCTCTGGCCCTGGCACTGGTCT
CCATGGCCTCAGCCACAGGGCTCCCTGGACCCCTCTTCCCTCCAGATCCCAGCCTCCTGCTGGGTAC
TGGTCTCATCTGGGCTAAAGTTTGAGACTGGCTCAAATCTCCCAAGCTGCTGCACGTGCTGAGTCCAGA
GGCAGTCACAGAGACCTCTGGCAGGGATCTAACACTGGGTTCTGGGTCTCAGGACTGAAGAGGAGGAG
TGGGGTCAAGAGATTCTCTGGCACCAAGTGCAGCATGCCACAAATCTTTAGGAATGGGACAGGTACCT
TCACTTGTGTGATTNN
CAGGAATGGCAGTAAATAAAAGTCTGCACTTGGTCAATTCTTCTCTCAGAGGAAGCCGAGTGTCACTAAAC
ACTATCCCCTCAGACTCCCTGTGTGAGGCCCTGCAGAGGCCCTGAATGCAAAAATGGGAAACCAAGGCACAGAG
GCTCTCTCTCTCTCTCCCCGATGTACCCCTCAAAAAAAATGCTAACCAAGTCTTCCATTAGC
CGGCTGAGTGAGGGAAAGCCAGCACTGCTGCCCTCTGGTAACTCACCTAAGGCTCTGGCCACCTCTGGCT
ATGGTAACCACACTGGGGCTTCTCCAAGCCCCGCTCTCCAGCAGTCCACCCAGGGCAGAGTCCAGAGCCAC
CACCCTGGGGTGGCTGTGGCCCCAGTCAGCTGCTCAGGACCTGCTATTTCAGGGAAAGAAGATTATGT
ATTATATGTTGCTATATTCTAGAGCACCTGTGTTTCTCTTCTAAGCCAGGGCTCTGTGATGACTTAT
GGGTGGGGAGGTGAAACCGGAACCTTCTATTGAAGGCAGTAAACTGTGCTAAATGCA

FIGURE 219

MSVMVVRKKVTRKWEKLPGRNTFCCDGRVMMARQKGIFYLTFLILGTCTLFFAFECRYLAV
QLSPAIPVFAAMLFLLFSMALLRTSFSDPGVIPRALPDEAAIFIEMEIEATNGAVPQGQRPPP
RIKNFQINNQIVKLKYCYTCKIFRPPRASHCSICDNCVERFDHHCPWVGNCVGKRNRYRFYL
FILSLSLLTIYVFAFNIVYVALKSLKIGFLETLKETPGTVLEVLCICFTLWSVVGLTGFHTF
LVALNQTTNEDIKGSWTGKNRVQNPYSHGNIVKNCCEVLCGPLPPSVLDRRGILPLEESGR
PPSTQETSSSLPQSPAPTEHLSNEMPEDSSTPEEMPPPEPPQEAEEAK

Putative transmembrane domains:

amino acids 36-55 (type II TM), 65-84, 188-208, 229-245

FIGURE 220

AAAACCTGTATTTTACAATGCAAATAGACAATNANCCTGGAGGTCTTGAATTAGGTAT
TATAGGGATGGTGGGGTTGATTTTNTTCCTGGAGGCTTGGCTTGGACTCTCNCTTCT
CCCACAGAGCNCTCGACCATCACTGCCCTGGTGGGAATTGTGTTGGAAAGAGGAACTA
CCGCTANTTCTACCTCTCATCCTTNTCTCTCCNCCTCACAAATCTATGTCTTCGCCTTCA
ACATCGT

卷之三

FIGURE 221

GTGTGTCTTCAGCAAAACAGTGGATTAAATCTCCTGCACAAGCTGAGAGCAACACAA
TCTATCAGGAAAGAAAGAAAGAAAAACCGAACCTGACAAAAAAGAAGAAAAAGAAGAAGA
AAAAAAAT**CATG**AAAACCATCCAGCCAAAATGCACAATTCTATCTCTGGCAATCTTCAC
GGGGCTGGCTGCTCTGTCTCTCCAAGGAGTGCCCGTGCAGCAGGGAGATGCCACCTTCC
CCAAAGCTATGGACAACGTGACGGTCCGGCAGGGGAGAGGCCACCCCTCAGGTGCACTATT
GACAACCGGGTCACCCGGGTGGCTGGCTAAACCGCAGCACCCTCTATGCTGGATGA
CAAGTGGTGCCTGGATCCTCGCGTGGCCTTCTGAGCAACACCCAAACGCACTACAGCATCG
AGATCCAGAACGTGGATGTGTATGACGAGGGCCCTACACCTGCTCGGTGCAGACAGACAAAC
CACCCAAAGACCTCTAGGGTCCACCTCATTGTCAAGTATCTCCAAAATTGTAGAGATTTC
TTCAGATATCTCCATTAATGAAGGAAACAATATTAGCCTCACCTGCATAGCAACTGGTAGAC
CAGAGCCTACGGTTACTTGGAGACACATCTCTCCAAAGCGGTTGGCTTGTGAGTGAAGAC
GAATACTGGAAATTCAAGGCATCACCCGGGAGCAGTCAGGGACTACGAGTGCAGTGCCTC
CAATGACGTGGCCGCCCGTGGTACGGAGAGTAAAGGTACCGTGAACATCCACCATACA
TTTCAGAAGCCAAGGGTACAGGTGCCCCGTGGACAAAAGGGACACTGCAGTGTGAAGCC
TCAGCAGTCCCCTCAGCAGAATTCCAGTGGTACAAGGATGACAAAAGACTGATTGAAGGAA
GAAAGGGTGAAAGTGGAAAACAGACCTTCTCTCAAAACTCATTTCAATGTCTTG
AACATGACTATGGAAACTACACTTGCCTGGCCTCCAACAAGCTGGCCACACCAATGCCAGC
ATCATGCTATTGGTCCAGCGCCGTAGCGAGGTGAGCAACGGCACGTCGAGGAGGGCAGG
CTCGTCTGGCTGCCTTCTGGTCTTGCACCTGCTCTCAAATTT**TGA**TGTGAGTGC
ACTTCCCCACCCGGAAAGGCTGCCACCACCAACACAACAGCAATGGCAACAC
CGACAGCAACCAATCAGATATACAAATGAAATTAGAAGAAACACAGCCTCATGGACAGA
AATTGAGGGAGGGAAACAAAGAATACTTGGGGAAAAGAGTTTAAAAAGAAATTGAA
AATTGCCTTGCAGATATTAGGTACAATGGAGTTCTTCTTCCAAACGGGAAGAACACAGC
ACACCCGGCTGGACCCACTGCAAGCTGCATCGTCAACCTCTTGGTGCAGTGTGGCAA
GGGCTCAGCCTCTGCCAACAGAGTGCCTTGGAACATTCTGGAGCTGCCATCCCA
AATTCAATCAGTCATAGAGACGAACAGAATGAGACCTCCGGCCAAGCGTGGCGTGC
GCACCTTGGTAGACTGTGCCACCACGGCGTGTGAAACGTGAAATAAAAGAGCAAA
AAAAA

FIGURE 222

MKTIQPKMHNSISWAIFTGLAALCLFQGVPRSGDATFPKAMDNVTVRQGESATLRCTIDNR
VTRVAWLNRSTILYAGNDKWCLDPRVVLLSNTQTQYSIEIQNVDVYDEGPYTCVQTDNHPK
TSRVHLIVQVSPKIVEISSDISINEGNNISLTCIATGRPEPTVTWRHISPKAvgFVSEDEYL
EIQGITREQSGDYECASNDVAAPVVRVKTVNYPYISEAKGTGVPVGQKGTLQCEASAV
PSAEFQWYKDDKRLIEGKKGVKVENRPFLSKLIFFNVSEHDYGNYTCVASNKLGHTNASIML
FGPGAVSEVSNGTTSRRAGCVWLLPLLVLHLLLKF

Signal peptide:

amino acids 1-28

FIGURE 223

GAAAAAAAATCATGAAAACCATCCAGCCAAAATGCACAATTCTATCTCTGGGCAATCTTC
ACGGGGCTGGCTGCTCTGTGTCTCTTCCAAGGAGTGCCCGTGCAGCGGAGATGCCACCTT
CCCCAAAGCTATGGACAACGTGACGGTCCGGCAGGGGGAGAGCGCCACCCCTCAGGTGCACTA
TTGACAACCGGGTCACCCGGGTGGCCTGGCTAAACCGCAGCACCATCCTCTATGCTGGGAAT
GACAAGTGGTGCCTGGATCCTCGCGTGGTCCTCTGAGCAACACCCAAACGCAGTACAGCAT
CGAGATCCAGAACGTGGATGTGTATGACGAGGGCCCTTACACCTGCTCGGTGCAGACAGACA
ACCACCCAAAGACCTCTAGGGTCCACCTCATTGTGCAAGTATCTCCAAAATTGTAGAGATT
TCTTCAGATATCTCCATTAATGAAGGAAACAATATTAGCCTCACCTGCATAGCAACTGGTAG
ACCAGAG

3
2
1
0
-1
-2
-3
-4
-5
-6
-7
-8
-9

FIGURE 224

ATGGCTGGTGACGGCGGGCGGGCAGGGGACCGGGGCCGGCGGGGGAGCAGCTGCGGGAGCCCTGA
ATCACCGCCTGGCCGACTCCACCATGAACGTCGCGCTGCAGGAGCTGGGAGCTGGCAGCAACGTTGGATTCCAG
AAGGGGACAAGACAGCTGTTAGGCTCACGCCACGCGAGCTGGAGCTGGCTTAGCAGGTGCTCTACTGCTGGCT
GCAC TGCTCTGGCTGCCTTGCGGCTAGGGTCCAGTACCAACAGAGACCCATCCCACAGCACCTGCCTTACA
GAGGGCTGCATTGAGTGGAAAAATCTGGAGCTGGACGGGGTGGAGCTGGGAGCCCTGTGAGGACTTTAC
CAGTTCTCTGTGGGGCTGGATTGGAGGAAACCCCTGGCGATGGCGTCTCGCTGGAACACACCTCAACAGC
CTCTGGGACCAAAACAGGCCATCTGAAGCACCTGTTGAAAACACCACTTCAACTCCAGCAGTGAAGCTGAG
CAGAAGACACAGCGCTTCTACCTATCTTGCTCACAGTGGAGCCATTGAGGAGCTGGGAGCCAGCCACTGAGA
GACCTCATTGAGAAGATTGGTGGTTGGAACATTACGGGGCCCTGGGACCAGGACAACATTATGGAGGTGAG
GCAGTAGCAGGGACCTACAGGGCCACCCATTCTCACCGTCTACATCAGTGGCAGCTAAGAGTTCAACAGC
AATGTTATCAGGTGGACCACTGGGCTCTTCTGCCCTCTGGGATTACTACTTAAACAGAACTGCCAATGAG
AAAGTGCTCACTGCCTATCTGGATTACATGGAGGAACCTGGGATGCTGCTGGGTGGCGGCCACCTCCACGAGG
GAGCAGATGCAGCAGGTGCTGGAGTTGGAGATAACAGCTGGCCAACATCACAGTGGCCACAGGACAGCGGCGCAG
GAGGAGAAGATCTACCACAAGATGAGCATTGGAGCTGAGCTGAGCCTGTTGGGTGAGTT
CTGTCTTCTTGTGTGTCACCATTGGAGTTGAGTGA
CAGGTGTCAGAGCTCATCAACCGCACGGAAACAGCATCCTGAACAATTACCTGATCTGAACCTGGTCAAAAG
ACAACCTCAAGCTGGACCGACGCTTGAGTCTGCACAAGAGAAGCTGCTGGAGACCCCTATGGCACTAAGAAG
TCCGTGTGCCGAGGTGGCAGACCTGCATCTCAACACGGATGACGCCCTGGCTTGGCTTGGGTCACTCTC
GTGAAGGCCACGTTGACCGGCAAAGAAATTGAGGAGGATGATCAGCGAAATCCGGACCGCATTGAG
GAGGCCCTGGACAGCTGGTTGGATGAGAAGACCCGCCAGGAGCAAGGAGAAAGCAGATGCCATCTAT
GATATGATTGGTTTCCAGACTTATCTGGAGCCAAAGAGCTGGATGATGTTTATGACGGGTACGAAATTCT
GAAGATTCTTCTTCCAAAACATGTTGAATTGAGTACAACCTCTGCTGCCAGGTTATGGCTGACAGCTCCGCAAG
CCTCCCAGCCGAGACCACTGGAGCATGACCCCCCAGACAGTGAATGCCACTACCTTCAACTAAGAATGAGATC
GTCCTCCCCCTGGCATCCTGCAGGCCCCCTCTATGCCGCACCCCAAGGCCCTGAACCTGGTGGCATC
GGTGTGGTCATGGCCATGAGTTGACGCTGCCCTTGATGACCAAGGGCGCAGTATGACAAAGAAGGGACCTG
CGGCCCTGGCAGAATGAGTCCCTGGCAGCCTTCCGGAACACAGGCCCTGCATGGAGGAACAGTACAATCAA
TACCAAGGTCAATGGGAGAGGCTAACCGCCAGACGCTGGGGAGAACATTACTGACAACGGGGGCTGAAG
GCTGCCCTAACATGCTTACAAAGCATGGCTGAGAAAGCATGGGGAGGAGCAGCAACTGCCAGCCGGCTCACC
AACCAACAGCTTCTCGTGGATTGCCCCAGGTGTTGCTCGGTCCGCACACCAAGAGAGCTCTCACGAGGGG
CTGGTGACGGACCCCCACAGCCCTGCCCGCTCCGCGTGGCACTCTCCAACTCCCGTGACTTCCCTGCGG
CACTCGGCTGCCCTGCGCTCCCCATGAACCCAGGGCAGCTGTGAGGTGTTGAGCTGGTAGACCTGGATCAGGGGA
GAAATGGCCAGCTGTCAACAGACCTGGGGAGCTCTCCTGACAAAGCTGTTGCTCTGGGTGGGAGGAAGCAA
ATGCAAGCTGGCTGGGTCTAGTCCCTCCCCCCCACAGGTGACATGAGTACAGACCCCTCTCAATCACCACATTG
TGCCTCTGCTTTGGGGTGCCTCCAGCAGAGCCCCCACCATTCACTGTGACATCTTCCGTGTCACCCT
GCCTGGAAGAGGTCTGGGTGGGGAGGCCAGTCCCATAGGAAGGAGTCTGCC

FIGURE 225

MNVALQELGAGSNVGFQKGTRQLLGSRTQLELVLAGASLLLALLLGCLVALGVQYHRDPSH
STCLTEACIRVAGKILESLDRGVSPCEDFYQFSCGGWIRRNPPLDGRSRWNTFNSLWDQNQA
ILKHLLENTTFNSSEAEQKTQRFYLSCLQVERIEELGAQPLRDLIEKIGGWNIITGPWDQDN
FMEVLKAVAGTYRATPFFTYYISADSKSSNSNVIQVDQSGFLPSRDYYLNRTANEKVLTAY
LDYMEELGMLLGGRPTSTREQMQQVLELEIQLANITVPQDQRRDEEKIYHKMSISELQALAP
SMDWLEFLSFLLSPLLESDSEPVVVYGMMDYLQQVSELINRTEPSILNNYLIWNLVQKTTSSL
DRRFESAQEKLLETLYGTTKSCVPRWQTCISNTDDALGFALGSLFVKATFDRQSKEIAEGMI
SEIRTAFFEEALGQLVWMDEKTRQAAKEKADAIYDMIGFPDFILEPKELDDVYDGYEISEDSF
FQNMLNLYNFSAKVMADQLRKPPSRDQWSMTPQTVNAYYLPTKNEIVFPAGILQAPFYARNH
PKALNFGGIGVVMGHELTHAFDDQGREYDKEGNLRPWWQNESLAAFRNHTACMEEQYNQYQV
NGERLNGRQTLGENITDNGGLKAAYNAYKAWLKHGEEQQLPAVGLTNHQLFFVGFAQVWCS
VRTPESSHEGLVTDPHSPARFRVLGTLNSRDFLRHFGCPVGSPMNPGQLCEVW

Type II Transmembrane domain:

amino acids 32-57

FIGURE 226

GGCCGGCCCTCGCCCTCCGACTCCGCCTCCCTCCCTCCGCCGCTCCCGGCCCTCCCTCCCTCCCTCC
CAGCTGTCCCGTTCGCGTCATGCCGAGCCTCCGCCGCCGCCAGAGCCCCCGTGTGCTCCCTCGGCTGCTGCTGCT
CGGCTCCCGCCGGCCGCCGCCGCCAGAGCCCCCGTGTGCTCCCTCGGCTGCTGCTGCTGCTGCTGCTGCTGCT
CGTTCGGGAGCGGCAGGTAGGTGGCGCCGGGGAGGCGCGGGGGAGTCGGGCTCGGGCTGGGGAGTCAGCGC
CAGCCCCGGAGGGGGCGCGGGCGCAGGTGGCTCGGCCGGCGGGGGAGGGTGGGCGGGGGAGAAGGG
GCGGTGCCTGGGACCCGGGACCCCGCGGGCAGCCCCGGGGCGGCACACGGCGCAGCTGGCAGCGGCCCTCCAGC
CAAGCCCCTCCCGCAGGCTGCACCTCGCGGGAAAGGTCTATGCCCTGGACGAGACGTGGCACCCGGACCTAGG
GGAGCCATTGGGTGATGCGCTGCGTGTGCGCTCGGAGGGCAGTGGGCTGCGTACCCAGGGGCCCTGG
CAGGGTCAGCTGCAAGAACATCAAACAGAGTGCACCCGGCTGTGGCAGCCGCGCAGCTGCCGGACA
CTGCTGCCAGACCTGCCCGAGGACTTCGTGGCGCTGTGACAGGGCGAGGTGCGCAGGGGTGGCACGAGCCCG
AGTCTCGCTGCGCTAGCCTCCGCTTCTATCCTACAGCGGCTGGACCGCCCTACAGGATCCGCTT
CTCAGACTCCAATGGCAGTGTCTGTTGAGCACCCCTGAGCCCCCACCAAGATGGCTGGTCTGTGGGTGTG
GCGGGCAGTGCCTCGGTTGTCTCGGCCCTCTAGGGCAGAACAGCTGCATGTGGACTTGTGACACTCACTCA
CCCTCAGGGGAGGTCTGGGGCCTCTCATCCGGCACCGGGCCCTGTCCCCAGAGACCTCAGTGCCTACCTGAC
TCTAGAAGGGCCCCCACCAGCAGGGCGTAGGGGCATACCCCTGCTCACTCAGTGCACACAGAGGACTCCTGCA
TTTTTGCTGCTCTCCGAGGCCTTGAGGACTAACCAAGGTTCCCTGAGGCTCCAGATTCTACACCAGGGCA
GCTACTGCGAGAACTTCAGGCCAATGTCTCAGCCCAGGAACCAGGCTTGCTGAGGTGCTGCCAACCTGACAGT
CCAGGAGATGGACTGGCTGGTGTGGAGCTGCAGATGGCCCTGGAGTGGCAGGCCAGGGCTGCGCAT
CAGTGGACACATTGCTGCCAGGAAGAGCTGCGACGCTCTGCAAAGTGTCCCTTGTGGGCTAATGCCCTGATCCC
AGTCCAAACGGGTGCTGCCGCTCAGCCAGCCTACTCTGCTAGGAAATGCCNCCCTGATCCTCCAGGTGCAATT
GGTAGGGACAACCAAGTGGAGGTGGGCTGACACTGGAAACCAAGCCTAGCGGAGGGATCACCCACTGTCCT
GTGCCACATGGCTGGCTATCCTCCCTGGGGCCAGGCCGTGGTATCTGCCCTGGCTGGGCTGGGGCGAGGGGC
TCATATGCTGCTGCAAGATGAGCTTCTCTGAAACGTGGCACCAGGACTTCCAGACGGAGAGCTCAGGGGCA
ACGTGGCTGCCCTGGGCTACTGTGGGCAATAGGCCCTGGGCTGGGGCTGGGGCTGGGGCTGGGCTGGGCT
CCCTGTGAAGAGCCAAGCAGCAGGGCACGCCCTGGCTTCTTGATACCCACTGTCACTGCACATGAAGTGT
GCTGGCTGGGCTGGCTCAGAACAGGCAGTGTCACTGCCACCTCTGGGCTCTGGGAACGCCAGGGC
TCGGCGGCTGCTGAAGGGATTCTATGGCTCAGAGGCCAGGGTGTGGTGAAGGACCTGGAGGCCAGTGTGCG
GCAACCTGGCAAAGGCAATGGCTTCCCTGATGATCACCAAGGTAGCCCAGAGGGAGCTCCGAGGGCAGCCT
CTCCTCCCAGGTGCACATAGCCAACCAATGTGAGGTGGCGACTGCCCTGGAGGCCGGGGGGAGGGGG
GCGGGCGCTGGGGCTCCGGATACAGCCTCTGCGCCGCGCTGTGGTGCTGGTCTCCGGGCTAGGCCCGC
CAAACCTGGTGGCTGGCGGCCAGAACACATGCTTCTTGAGGGGAGCAGGCCAGGGGGAGCTCCGAGGGCAGC
TCGCTGGGCCCAACTACGACCCGCTCTGCACTCTGCACCTGCCAGAGACGAACGGTATGTGACCCGGT
GGTGTGCCACCGCCAGCTGCCACACCCGGTGCAGGCTCCGACCGAGTGTGCCCTGTTGCCCTGGCTGCTA
TTTGATGGTGAACGGAGCTGGCGGAGCAGGGTACCGGGTGGCACCCGGTGTGCCCTGGCTTAATTAA
GTGTGCTGCTGCAACCTGCAAGCAGGGGGCACTGGAGAGGTGCACTGTGAGAAGGTGCACTGTCCCCGGCTGG
CTGTGCCACGGCTGTGCGTCAACCCACCGACTGTCACACAGTGTCCAGGTGAGGCCACCCAGCTGG
GGACCCCATGCAAGGCTGATGGGCCGGCTGCCGTTTGCTGGCAGTGGTCCAGAGACTCAGAGCTGGCA
CCCTCAGTGCCTGGGCTGGAGAGATGAGCTGTATCACCTGCAAGATGTGGGTAAGTGGGAGCAGAGGCTGT
GTGAGGTGGTACTGGGAGCTGGCTGGAGTAGGGAGACCTTCCAGGGAGGGCTCCCTGAAGAACGCTGAAGGTCA
CTGTGCTCCAGTGCCTCTGGGGACACTCAGTGTCTGTCTGTGTCACCTGCAAGGGGTGCCCTACTGTGAGC
GGATGACTGTTACTGCCACTGTCTGTGCTGGCTGGGAGAGGAGACTGCACTGCTGTTCCCCTGCACGCC
GGCGCGTAAGTGAGGGAGTCCAGGGTCAAGCAGCTGCACTGCCCTGCGCTCACAGTGCACCTGCA
GGAAGGGAGCACTCACTGTGTGCAAGAACAGTGCAGGCCCTGCGCTCACAGTGCACCTGCA
ACCTGGTGAATTGTTATTGACCTTTCTTACAAATGAGATTCTGAAGCTCAGAGAAATTAAAGCAACGAG
ATGAAGGTCAACCCAGCTGTGTGCACTGACCTGTTAGAAAATACTGGCCTTCTGGGACCAAGGCAGGGATGCTT
TGCCCTGCCCTCTATGCCCTCTGTGCCCTCCACTCCCTCTCCCTCCAAACATCCCTCCCTGTCTCC
AGCAGCCCCAGAGACCAACTGATCCAGAGACTGGAGAAAGAACGCCAGGGCTTAAAGGAGCAGGCCAGGG
AAGTGAACCAAGAGGATGGGGCTGAGCTGGGGAGGGGTGGCATCGAGGACCTTCTGCATTCTCCTGTGG
CCCAGTGCCTTGCTCTGTGCTCTACTCCCACCCCAACTACCTCTGGGACCAAGCTCCACAAGGG
GAGAGGAGCTGGGCCAGACCGAGGTCACAGCACTCCAAGTCTGCCCTGCCACCCCTGGCCTGTCTGGAA
GCCCAACCCCTTCTCTGTACATAATGTCAGTGGCTGTTGGATTTTAATTATCTCACTCAGCACCAAG
GGCCCGGACACTCCACCTCTGCGCCCTGAGCTGAGCAGAGTCATTATGGAGAGTTGTATTAAAC
ATTCTTTTCAGTCTTGGCATGAGGTGGCTTTGTGGCAGGAACCTGAGTGGGCTGGGAGAAGGG
GCNGAGAGTAGGAGGTGAGAGAGAGGAGCTGTGACACTTGGGGAGCTGAAAGAGACCTGGAGAGGCAGAGG
CGTGGCNNTGGCTGGCATNCCTGGGTTCCGAGAGGGGCTGGGATGGTCTTGAGATGGCTAGAGACTCAAG
AATTAGGGAAAGTAGAAGCAGGATTGACTCAAGTTAGTTCCCACATCGCTGCCCTGTTGACTTCATG
TTGAGTTGCTCCAGAGAGAGAATCAAAGGTGTCACCAGCCCTCTCCCTCCCTCCCT
TTCCCTCCCTCCCTCCCTCCCTCCCTCC

FIGURE 227

GGCGAGCGGGGGTGTGCGCGCGGCCGTGATGGCTGGTACGGCGGGCCGGCAGGGGA
CCGGGGCCGCGGCCGGAGCGGGCCAGCTGCCGGAGCCCTGAATACCGCCTGGCCCGAC
TCCACCATGAACGTGCGCTGCAGGAGCTGGAGCTGGCAGCAACGTGGATTCCAGAAGGG
GACAAGACAGCTGTTAGGCTCACGCACGCAGCTGGAGCTGGTCTTAGCAGGTGCCTCTAC
TGCTGGCTGCACTGCTCTGGCTGCCTGTGCCCTAGGGTCCAGTACCAACAGAGACCCA
TCCCACAGCACCTGCCTTACAGAGGCCTGCATTGAGTGGCTGGAAAAATCCTGGAGTCCCT
GGACCGAGGGGTGAGCCCCGTGAGGACTTTACCAAGTTCTCCTGTGGGGCTGGATTCGGA
GGAACCCCTGCCGATGGCGTTCTCGCTGGAACACCTAACAGCCTCTGGACCAAAAC
CAGGCCATACTGAAGCACCTGTTGAAAACACCACCTCAACTCCAGCAGTGAAGCTGAGCA
GAAGACACAGCGCTTCTACCTATCTGCCTACAGGTGGAGCGCATTGAGGAGCTGGAGCCC
AGCCACTGAGAGACCTCATTGAGAAGATTGGTGGTGGAACATTACGGGCCCTGGGACCAAG
GACAACTTATGGAGGTGTTGAAGGCAGTAGCAGGGACCTACAGGGCCACCCATTCTCAC
CGTCTACATCAGTGCCGACTCTAACAGAGTTCCAACAGCAATGTTATCCAGGTGGACCAAGTCTG
GGCTTTCTGCCCTCTGGGATTACTACTTAAACAGAACTGCCAATGAGAAAGTAAGGAAC
ATCTTCCGAACCCCCATCCCTACCCCTGGCTGAGCTGGCTGATCCCTGTTGACTTTCCCT
TTGCCAAGGGTCAGAGCAGGGAAAGGTGAGCCTACCTGTACCTAGTGAACAAACTGCCCT
CCTTCTTCTTCTTCTCCCTCCCTCCCTTCTCCCTTCTCCCTTCTCCCTTCC
TCTTATTCTCTAGTAGGTTCATAGACACCTACTGTGTGCCAGGTCCAGTGGGGAAATTG
GAGATATAAGTTCCGAGCCATTGCCACAGGAAGCGTTAGTGTGATGGGTCATGGACCT
AGATAGGCTGATAACAAAGCTCACAAAGAGGGCCTGAGGATTAGGAGAGACTTATGGAGCC
AGCAAAGTCTCCTGAAGAGATTGCATTGAGCCAGGTCTGTAG

FIGURE 228

ATGCCTACTACCTCCA ACTAAGAATGAGATCGTCTTCCCCGCTGGCATCCTGCAGGCC
TTCTATGCCCGCAACCACCCCAAGGCCCTGAACCTCGGTGGCATCGGTGTGGCATGGCCA
TGAGTTGACGCATGCCTTGATGACCAAGGGCGCGAGTATGACAAAGAACCTGCGGC
CCTGGTGGCAGAATGAGTCCCTGGCAGCCTCCGGAACACAGGCCTGCATGGAGGAACAG
TACAATCAATACCAGGTCAATGGGAGAGGGCTCAACGGCCGCCAGACGCTGGGGAGAACAT
TGCTGACAACGGGGGCTGAAGGCTGCCTACAATGCTTACAAAGCATGGCTGAGAAAGCATG
GGGAGGAGCAGCAACTGCCAGCCGTGGGCTCACCAACCACAGCTCTCAGAGGGCTGGTGACCGACCC
GCCAGGTGTGGTGCCTGGTCCGACACCAGAGAGCTCTCACGAGGGCTGGTGACCGACCC
CCACAGCCCTGCCCGCTTCCCGTGTGGCACTCTCCAACCTCCGTGACTTCCTGCGGC
ACTTCGGCTGCCCTGTCGGCTCCCCATGAACCCAGGGCAGCTGTGTGAGGTGTGGTAGACC
TGGATCAGGGGAGAAATGCCAGCTGTCACCAGACCTGGGAGCTCTCCTGACAAAGCTGT
TTGCTCTGGGTTGGAGGAAGCAAATGCAAGCTGGCTGGTCTAGTCCCTCCCCCACA
GGTGACATGAGTACAGACCCCTCAATCACCAATTGTGCCCTGCTTGGGGTGCCCT
GCCTCCAGCAGAGCCCCCACCATTCACTGTGACATCTTCCGTGCACCCGTGGAAGAG
GTCTGGGTGGGAGGCCAGTTCCCATAGGAAGGAGTCTGCCTCTGTCCCCAGGCTCACT
CAGCCTGGCGGCCATGGGCCTGCCGTGCCACTGTGACCCACAGGCCTGGGTGGTG
TACCTCCTGGACTTCTCCCCAGGCTCACTCAGTGCACCTAGGGTGGACTCAGCTGT
TGGCTCACCCCTCACGGCTACCCCCACCTCACCCGTGCTGCCACTGCTCCAGTG
CTGCTGCTGACCTCACTGACAGCTCTAGTGGAAAGCCAAGGGCTCTGAAAGCCTCTGC
TGCCCACTGTTCCCTGGCTGAGAGGGAAAGTGCATATGTGTAGCGGGTACTGGTTCTGT
GTCTTAGGGCACAAGCCTAGCAAATGATTGATTCTCCCTGGACAAAGCAGGAAAGCAGATA
GAGCAGGGAAAAGGAAGAACAGAGTTATTTACAGAAAAGAGGGTGGAGGGTGTGGTCT
TGGCCCTTATAGGACC

FIGURE 229

CCCACGCGTCCGAGCCGCCGAGAATTAGACACACTCCGGACGCCAAAAGCAACCGAGA
GGAGGGGAGGCAAAAACACCGAAAAACAAAAGAGAGAAACAAACACCCAACTGGGTGG
GGGGAAAGAAAGAAAAGAAACCCACCCACCAAAAAA
AAAAAAAAAAATCCTGTGGCGCGCCTGGTCCCGGAAGACTCGCCAGCACCAGGGGG
TGGGGAGTGCAGCTGAAAGCTGCTGGAGAGTGAGCAGCCTAGCAGGGATGGAC**ATGATG**
CTGTTGGTGAGGGTCTGCTCGAACCAAGTGGCTGGAGAGTGAGCAGCCTAGCAGGGATGGAC**ATGATG**
CTGCCTGCTACCCCTCGCTCCCGCTGGACAGAGTGAGCTGGACTTCCCTGGCGGCCGTGG
ACAACATGATGGTCAGAAAAGGGACACGGCGGTGCTTAGGTGTTATTGGAAGATGGAGCT
TCAAAGGGTGCCTGGCTAACCGGTCAAGTATTATTGCGGGAGGTGATAAGTGGTCAGT
GGATCCTCGAGTTCAATTCAACATTGAATAAAGGGACTACAGCCTCCAGATAACAGAATG
TAGATGTGACAGATGATGGCCCATAACAGTGTCTGTTAGACTCAACATAACCCAGAAC
ATGCAGGTGCATCTAACTGTGCAAGTTCTCTTAAGATATATGACATCTCAAATGATATGAC
CGTCAATGAAGGAACCAACGTCACTCTTACTTGGCCACTGGGAAACCAGAGCCTCCA
TTCTTGGCGACACATCTCCCATCAGCAAACCATTTGAAAATGGACAATATTGGACATT
TATGGAATTACAAGGGACCAGGCTGGGAATATGAATGCAGTGCAGGAAATGCTGTGTCATT
CCCAGATGTGAGGAAAGTAAAAGTTGTTGTCACCTTGCTCCTACTATTAGGAAATTAAAT
CTGGCACCGTGACCCCCGGACGCAGTGGCCTGATAAGATGTGAAGGGTGCAGGTGTGCC
CCAGCCTTGAATGGTACAAAGGAGAGAAGAAGCTCTCAATGGCAACAAGGAATTATT
TCAAATTTAGCACAAGATCCATTCTACTGTTACCAACGTGACACAGGAGCACTCGGCA
ATTATACTGTGGCTGCCAACAGCTAGGCACAACCAATGCAGGCTGCCCTTAACCC
CCAAGTACAGCCCAGTATGGAATTACGGGAGCGCTGATGTTCTCTGCTGGTACCT
TGTGTTGACACTGTCCTTTACCAGCATATTCTACCTGAAGAATGCCATTCTACAA**AAA**
TTCAAAGACCCATAAAAGGCTTTAAGGATTCTCTGAAAGTGTGATGGCTGGATCAAATCT
GGTACAGTTGTTAAAAGCAGCGTGGATATAATCAGCAGTGCTTACATGGGATGATGCC
TTCTGTAGAATTGCTCATTATGTAATACTTAATTCTACTCTTTGATTAGCTACATTA
CCTTGTGAAGCAGTACACATTGCTTTTAAGACGTGAAAGCTGAAATTACTTTAG
AGGATATTAAATTGTGATTTCATGTTGTAATCTACAACCTTCAAAAGCATTCACTGATGGT
CTGCTAGGTTGCAGGCTGTAGTTACAAAAGCAATTGCAAGTGAATATGTGATTCTTAA
GGCTGCAATACAAGCATTCAAGTCCCTGTTCAATAAGAGTCAATCCACATTACAAAGATG
CATTCTTCTTTGATAAAAAGCAAATAATTGCTTCAGATTATTCTCAAATA
TAACACATATCTAGATTCTGCTGCATATTCAAGGTTCAAGGAATGAGCCTTGAAT
ATAACTGGCTGTGCAGCTGCTCTTCTGTAAGTTCAAGCATGGGTGTGCCCTCATAC
AATAATATTCTCTTGTCTCCAACTAATATAAAATGTTGCTAAATCTACAATTG
AAAGTAAAATAAACCAAGAGTGATCAAGTTAAACCATACACTATCTCAAGTAACGAAGGAGC
TATTGGACTGTAATACTCTTCTGCACTGACAATGGGTTGAGAATTGCCCCACACT
AACTCAGTTCTGTGATGAGAGACAATTAAACAGTATAGTAAATACCATATGATTTC
TTTAGTTGAGCTAAATGTTAGATCCACCGTGGAAATCATTCCCTTAAATGACAGCACA
GTCCACTCAAAGGATTGCCTAGCAATACAGCATCTTCTTCACTAGTCCAAGCCAAAA
TTTAAGATGATTGTCAGAAAGGGACAAAGTCCTATCACCTAATATTACAAGAGTTGGTA
AGCGCTCATCTTGTGAGACTATAATGGCAGAAACCAAGGGTGGTATGACAGTCGAC
TATGGATATGGACGTTCCAGAGACTATAATGGCAGAAACCAAGGGTGGTATGACCGCTACTC
AGGAGGAAATTACAGAGACAATTATGACAACAGTGAATGAGACATGCACATAATAGATA
CAAGGAATAATTCTGATCCAGGATCGCCTTCAAATGGCTGTATTATAAAGGTTTGG
AGCTGCACTGAAGCATTCTATTAGTATATCAACCTTGTAAATTGACGCTATTAA
AGGTAGCTGAAGACCTTTAGACAGTCCATCTTTAAATTCTGCCTATTAA
AGACAAATTATGGGACGTTGTCAAAAAAAAAAAAAAA

FIGURE 230

MMLLVQGACCSNQWLAAVLLSLCCLLPSCLPAGQSVDFPWAAVDNMMVRKGDTAVLRCYLED
GASKGAWLNRSSIIFAGGDKWSVDPRVSISTLNKRDYSLQIQNVDVTDDGPYTC CSVQTQHTP
RTMQVHLTVQVPPKIYDISNDMTVNEGTNVTLTCLATGKPEPSISWRHISPSAKPFENGQYL
DIYGITRDQAGEYECSAENAVSF PDVRKVKVVNFAPTIQEIKSGTVTPGRSGLIRCEGAGV
PPPAFEWYKGEKKLFNGQQGIIIQNFSTRSILT VTNVTQEHFGNYTCVAANKLGTNASLPL
NPPSTAQYGITGSADVLFSCWYLVLTLSFTSIFYLKNAILQ

Important features of the protein:

Signal peptide:

amino acids 1-31

Transmembrane domain:

amino acids 326-345

N-glycosylation sites.

amino acids 71-75, 153-157, 273-277, 284-288, 292-296, 305-309

Casein kinase II phosphorylation site.

amino acids 147-151, 208-212, 224-228

Tyrosine kinase phosphorylation site.

amino acids 178-186

N-myristoylation sites.

amino acids 7-13, 63-70, 67-73, 151-157, 239-245, 291-297,
302-308, 319-325

Myelin P0 protein:

amino acids 92-121

FIGURE 231

AGTGGTTCGATGGGAAGGATCTTCTCCAAGTGGTCCTCTGAGGGGAGCATTCTGCTGG
CTCCAGGACTTGGCCATCTATAAAGCTTGGCAATGAGAAATAAGAAAATTCTAAGGAGGA
CGAGCTCTGAGTGAGACCCAACAAGCTGCTTACCAAATTGCAATGGAGCCTTCGAAA
TCAATGTTCCAAGGCCAAGAGGAGAAATGGGTGAACCTCTCCCTAGCTGTGGTGGTCATC
TACCTGATCCTGCTCACCGCTGGCGCTGGCTGCTGGTCCAAGTTCTGAATCTGCAGGC
GGGGCTCCGGTCCTGGAGATGTATTCCTCAATGACACTCTGGCGCTGAGGACAGCCGT
CCTTCTCCTGCTGCAGTCAGCACACCCCTGGAGAACACCTGGCTCAGGGTCATCGAGGCTG
CAAGTCCTGCAGGCCAACTCACCTGGTCCCGTCAGCCATGAGCACTTGCTGCAGGGT
AGACAACTTCACTCAGAACCCAGGGATGTTAGAATCAAAGGTGAACAAGGCCAGGTC
TTCAAGGTACAAGGGGCCATGGCATGCCTGGTCCCCCTGGCCGCCGGACCACCTGCT
GAGAAGGGAGCAAGGGCTATGGGACGAGATGGAGCAACAGGCCCTGGGACCCAGG
CCCACCGGGAGTCAAGGGAGAGGCCCTCCAAGGACCCAGGGTGCCTCAGGGAAAGCAAG
GAGCCACTGGCACCCAGGACCCCAAGGAGAGAAGGGCAGCAAAGGCGATGGGCTCATT
GGCCAAAAGGGAAACTGGAACTAAGGGAGAGAACGGAGACCTGGTCTCCAGGAAGCAA
AGGGGACAGGGCATGAAAGGAGATGCAGGGTATGGGCCTCCTGGAGGCCAGGGAGTA
AAGGTGACTTCGGGAGGCCAGGCCACCAGGTTGGCTGGTTTCTGGAGCTAAAGGAGAT
CAAGGACAACCTGGACTGCAGGGTGTCCGGGCCCTGGTGAGTGGACACCCAGGTGC
CAAGGGTGAGCCTGGCAGTGCTGGCTCCCTGGCGAGCAGGACTTCCAGGGAGGCCGG
GTCCAGGAGCCACAGGCCCTGAAAGGAAGCAAAGGGACACAGGACTTCAAGGACAGCAAGGA
AGAAAAGGAGAATCAGGAGTTCCAGGCCCTGCAGGTGTGAAGGGAGAACAGGGAGGCCAGG
GCTGGCAGGTCCAAGGGAGCCCTGGACAAGCTGGCAGAAGGGAGACCAGGGAGTGAAG
GATCTTCTGGGAGCAAGGAGTAAAGGGAGAAAAGGTGAAAGAGGTGAAAACACTCAGTGTCC
GTCAGGATTGTCGGCAGTAGTAACCGAGGCCGGCTGAAGTTACTACAGTGGTACCTGGG
GACAATTGCGATGACGAGTGGCAAATTCTGATGCCATTGCTTCTGCCGATGCTGGTT
ACTCCAAAGGAAGGGCCCTGTACAAAGTGGAGCTGGCACTGGCAGATCTGGCTGGATAAT
GTTCAAGTGTGGGAGCACGGAGAGTACCCCTGTGGAGCTGCACCAAGAATAGCTGGGCCATCA
TGACTGCAGCCACGAGGAGGACGCAGGCGTGGAGTGCAGCGTCTTGACCCGAAACCCTTCA
CTTCTCTGCTCCGAGGTGTCTCGGGCTCATATGTGGGAAGGCAGAGGATCTTGAGGAGT
TCCCTGGGACAACGTGAGCAGCCTGGAGAGGGGCCATTAATAAGCTCAACATCATTGA

FIGURE 232

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA68886

><subunit 1 of 1, 520 aa, 1 stop

><MW: 52658, pI: 9.16, NX(S/T): 3

MRNKKILKEDELLSETQQAAFHQIAMEPFEINVPKPKRRNGVNFSLAVVVIYLILLTAGAGL
LKVQVLNLQARLRVLEMYFLNDTLAAEDSPSFSLQSAPGEHLAQGASRLQVLQAQLTWVR
VSHEHLLQRVDNFTQNPGMFRIKGEQGAPGLQGHKGAMGMPGAPGPPGPPAEKGAKGAMGRD
GATGPGSPQGPPGVKEAGLQGPQGAPGKQGATGTPGPQGEKGSKGDGGLIGPKGETGKGE
KGDLGLPGSKGDRGMKGDAVGVMGPPGAQGSKGDFGRPGPPLAGFPGAKGDQGQPGLQGVPG
PPGAVGHPGAKGEPEGSAGSPGRAGLPGSPGSPGATGLKGSKGDTGLQQGQQGRKGESGVPGPA
GVKGEQGSPGLAGPKGAPGQAGQKGDQGVKGSSGEQGVKGEKGERGENSVSRIVGSSNRGR
AEVYYSGTWGTICDDEWQNSDAIVFCRMLGYSKRALYKVGAGTGQIWLDNVQCRGTESTLW
SCTKNSWGHHDSCHEEDAGVECSV

Transmembrane domain:

amino acids 47-66 (type II)

N-glycosylation sites.

amino acids 43-47, 83-87, 136-140

Tyrosine kinase phosphorylation site.

amino acids 432-440

N-myristoylation sites.

amino acids 41-47, 178-184, 253-259, 274-280, 340-346, 346-352,
400-406, 441-447, 475-481, 490-496, 515-521

Amidation site.

amino acids 360-364

Leucine zipper pattern.

amino acids 56-78

Speract receptor repeat

amino acids 422-471, 488-519

C1q domain proteins.

amino acids 151-184, 301-334, 316-349

FIGURE 233

CCACCGCGTCCGAAGGCAGACAAAGGTTCATTTGTAAGAAGCTCCTCCAGCACCTCCTCT
CTTCTCCTTTGCCCAAACTCACCCAGTGAGTGTGAGCATTAAAGAACATCCTCTGCCAAG
ACCAAAAGGAAAGAAGAAAAAGGGCCAAAGCCAAATGAAACTGATGGTACTTGTTCAC
CATTGGGCTAACTTGCTGCTAGGAGTTCAAGCCATGCCCTGCAAATGCCCTCTTGTACAC
GAAAGATACTAAAGATCACAACTGTACAAACCTCCGGAAAGGAGTAGCTGACCTGACACAG
ATTGATGTCAATGTCCAGGATCATTCTGGATGGGAAGGGATGTGAGATGATCTGTTACTG
CAACTTCAGCGAATTGCTCTGCTGCCAAAGACGTTTCTTGGACCAAAGATCTCTTCG
TGATTCCCTGCAACAATCAATGAGAATCTTCATGTATTCTGGAGAACACCATTCTGATTTC
CCACAAACTGCACTACATCAGTATAACTGCATTCTAGTTCTATATAGTGCATAGAGCAT
AGATTCTATAAATTCTTACTTGTCTAAGACAAGTAAATCTGTGTTAAACAAGTAGTAATAAA
AGTTAATTCAATCTAAAAAAAAAAAAAA

FIGURE 234

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52758
<subunit 1 of 1, 98 aa, 1 stop
<MW: 11081, pI: 6.68, NX(S/T): 1
MKLMVLVFTIGLTLGGVQAMPANRLSCYRKILKDHNCNLPEGVADLTQIDVNVQDHFWDG
KGCEMICYCNFSELLCCPKDVFVFGPKISFVIPCNNQ
```

Important features:

Signal peptide:

amino acids 1-20

N-glycosylation site.

amino acids 72-76

Tyrosine kinase phosphorylation site.

amino acids 63-71

FIGURE 235

CCCCACGCGTCCGCGGACGCGTGGGCTGGACCCCAGGTCTGGAGCGAATTCCAGCCTGCAGGG
CTGATAAGCGAGGCATTAGTGAGATTGAGAGAGACTTACCCCGCCGTGGTGGAGGGC
GCGCAGTAGAGCAGCAGCACAGGCGCGGGTCCCAGGCGGCTCTGCTCGCGCCGAG**ATG**
TGGAAATCTCCTTCACGAAACCGACTCGGCTGTGGCCACCGCGCGCCCGCCTGGCTGTG
CGCTGGGGCGCTGGTGTGGCTGGCGGGCTTCTTCTCCTCGGCTTCCTCTGGGTGGTTA
TAAAATCCTCAAATGAAGCTACTAACATTACTCAAAGCATAATATGAAAGCATTGGAT
GAATTGAAAGCTGAGAACATCAAGAAGTTCTTACATAATTTACACAGATACCACTTAGC
AGGAACAGAACAAAACCTTCAGCTGCAAAGCAAATTCAATCCCAGTGGAAAGAATTGGCC
TGGATTCTGTTGAGCTAGCTCATTATGATGTCCTGTTGTCACCCAAATAAGACTCATCCC
AACTACATCTCAATAATTAAATGAAGATGAAATGAGATTTCACACATCATTATTGAACC
ACCTCCTCCAGGGATATGAAAATGTTGGATATTGTACCACTTCAGTGCTTCTCCTC
AAGGAATGCCAGAGGGCGATCTAGTGTATGTTACTATGCACGAACGTGAAGACTTCTTAAA
TTGGAACGGGACATGAAAATCAATTGCTCTGGAAAATTGTAATTGCCAGATATGGGAAAGT
TTTCAGAGGAAATAAGGTTAAAATGCCAGCTGGCAGGGCCAAAGGAGTCATTCTCTACT
CCGACCCCTGCTGACTACTTGCTCTGGGTGAAGTCCTATCCAGACGGTTGGAATCTCCT
GGAGGGTGGTGTCCAGCGTGGAAATATCCTAAATCTGAATGGTGCAGGAGACCCCTCACACC
AGGTTACCCAGCAAATGAATATGCTTATAGGCGTGAATTGCAGAGGCTGTTGGTCTTCAA
GTATTCCCTGTTCATCCAATTGGATACTATGATGCACAGAACGCTCTAGAAAAAATGGGTGGC
TCAGCACCAACCAGATAGCAGCTGGAGAGGAAGTCTCAAAGTGCCTACAATGTTGGACCTGG
CTTTACTGGAAACTTTCTACACAAAAAGTCAAGATGCACATCCACTTACCAATGAAGTGA
CGAGAATTACAATGTGATAGGTACTCTCAGAGGAGCAGTGGAACCAAGACAGATATGTCATT
CTGGGAGGTCACCGGGACTCATGGGTGTTGGTGGTATTGACCCCTCAGAGTGGAGCAGCTGT
TGTTCATGAAATTGAGGAGCTTGGAACACTGAAAAGGAAGGGTGGAGACCTAGAAGAA
CAATTGTTGCAAGCTGGGATGCAGAAGAATTGGTCTTGGTTCTACTGAGTGGCA
GAGGAGAATTCAAGACTCCTCAAGAGCGTGGCGTGGCTTATATTAAATGCTGACTCATCTAT
AGAAGGAAACTACACTCTGAGAGTTGATTGACACCCTGATGTACAGCTGGTACACAACC
TAACAAAAGAGCTGAAAAGCCCTGATGAAGGCTTGAAGGCAAATCTTTATGAAAGTTGG
ACTAAAAAAAGTCCTCCCCAGAGTTGAGCTGGCATGCCAGGATAAGCAAATTGGGATCTGG
AAATGATTGAGGTTCTCCAACGACTTGAATTGCTTCAGGCAGAGCACGGTATACTA
AAAATTGGGAAACAAACAAATTCAAGCGGCTATCCACTGTATCACAGTGTCTATGAAACATAT
CAGTTGGGGAAAAGTTTATGATCCAATGTTAAATATCACCTCACTGTGGCCAGGTTCG
AGGAGGGATGGTGTGAGCTAGCCAATTCCATAGTGCTCCCTTGATTGTCGAGATTATG
CTGTAGTTTAAGAAAGTATGCTGACAAAATCTACAGTATTCTATGAAACATCCACAGGAA
ATGAAGACATACAGTGTATCATTGATTCACTTTCTGCAGTAAAGAATTTCACAGAAAT
TGCTTCCAAGTTCACTGAGAGACTCCAGGACTTGACAAAAGCAACCCAAATAGTATTAAAGAA
TGATGAATGATCAACTCATGTTCTGGAAAGAGCATTATTGATCCATTAGGGTTACCAAG
AGGCCTTTTATAGGCATGTCATCTATGCTCCAAGCAGCCACAACAAGTATGCAGGGAGTC
ATTCCCAGGAATTATGATGCTCTGTTGATATTGAAAGCAAAGTGGACCCTTCCAAGGCCT
GGGGAGAAGTGAAGAGACAGATTATGTTGCAGCCTCACAGTGCAGGCAGCTGCAGAGACT
TTGAGTGAAGTAGCC**TAA**AGAGGATTAGAGAATCCGTATTGAATTGTTGAGGTTATGTCA
CTCAGAAAGAATCGTAATGGGTATATTGATAAAATTAAAATTGGTATATTGAAATAAAAGT
TGAATATTATATAA

FIGURE 236

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA52756
><subunit 1 of 1, 750 aa, 1 stop
><MW: 84305, pI: 6.93, NX(S/T): 10
MWNLHETDSAVATARRPRWLCA GALVLAGGFFLLGFLFGWFIKSSNEATNITPKHNMKAFL
DELKAENIKKFLHNFTQIPLLAGTEQNFQLAKQIQSQWKEFGLDSVELAHYDVLLSYPNKTH
PNYISIINEDGNEIFNTSLFEP PPPGYENVSDIVPPFSAFSPQGMPEGDLVYVNYARTEDFF
KLERDMKINC SGKIVIARYGKVFRGNVKVNAQLAGAKGVILYSDPADYFAPGVKSYPDGWNL
PGGGVQRGNILNLNGAGDPLTPGYPANEYAYRRGIAEAVGLPSIPVHPIGYYDAQKLLEKMG
GSAPPDSSWRGSLKVPYNVPGFTGNFSTQKVKMHIHSTNEVTRIYNVIGTLRGAVEPDRYV
ILGGHRDSWVFGGIDPQSGAAVVHEIVRSFGTLKKEGWRPRRTILFASWDAEEFGLLGSTEW
AEENSRLLQERGVAYINADSSIEGNYTLRVDCPLMYSLVHNLTKE LKSPDEGFEGKSLYES
WTKKSPSPEFSGMPRI SKLGS GND FEVFFQRLGIASGRARYTKNWETNKFSGYPLYHSVYET
YELVEKFYDPMFKYHLTVAQVRGGMV FELANSIVLPFDCRDYAVVLRKYADKIYSISMKHPQ
EMKTYSVSFDSLFSAVKNFTEIASKFSERLQDFDKSNPIVLRMMNDQLMFLERA FIDPLGLP
DRPFYRHVIYAPSSHNKYAGESFPGIYDALFDIESKVDP SKAWGEVKRQIYVAAFTVQAAAETLSEVA

Signal sequence:

amino acids 1-40

N-glycosylation sites.

amino acids 76-80, 121-125, 140-144, 153-157, 195-199, 336-340,
459-463, 476-480, 638-642

Tyrosine kinase phosphorylation sites.

amino acids 363-372, 605-613, 606-613, 617-626

N-myristoylation sites.

amino acids 85-91, 168-174, 252-258, 256-262, 282-288, 335-341,
360-366, 427-433, 529-535, 707-713